U.S. Department
of Transportation
**National Highway
Traffic Safety
Administration**

www.nhtsa.dot.gov
nhtsa
people saving people

DOT HS 810 593

April 2006

The 100-Car Naturalistic Driving Study

Phase II – Results of the 100-Car Field Experiment

This document is available to the public from the National Technical Information Service, Springfield, Virginia 22161

This publication is distributed by the U.S. Department of Transportation, National Highway Traffic Safety Administration, in the interest of information exchange. The opinions, findings and conclusions expressed in this publication are those of the author(s) and not necessarily those of the Department of Transportation or the National Highway Traffic Safety Administration. The United States Government assumes no liability for its content or use thereof. If trade or manufacturer's names or products are mentioned, it is because they are considered essential to the object of the publication and should not be construed as an endorsement. The United States Government does not endorse products or manufacturers.

Technical Report Documentation Page

1. Report No. DOT HS 810 593	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle The 100-Car Naturalistic Driving Study, Phase II – Results of the 100-Car Field Experiment		5. Report Date April 2006 6. Performing Organization Code
7. Authors Dingus, T. A., Klauer, S.G., Neale, V. L., Petersen, A., Lee, S. E., Sudweeks, J., Perez, M. A., Hankey, J., Ramsey, D., Gupta, S., Bucher, C., Doerzaph, Z. R., Jermeland, J., and Knipling, R.R.		8. Performing Organization Report No.
9. Performing Organization Name and Address Virginia Tech Transportation Institute 3500 Transportation Research Plaza (0536) Blacksburg, Virginia 24061		10. Work Unit No. (TRAIS) 11. Contract or Grant No. DTNH22-00-C-07007
12. Sponsoring Agency Name and Address National Highway Traffic Safety Admin. (NHTSA) NPO-113, Room 6220 Office of Advanced Safety Research 400 Seventh Street SW, Room 6220 Washington, DC 20590		13. Type of Report and Period Covered Interim Report August 2001 to March 2005 14. Sponsoring Agency Code NHTSA NPO-113
15. Supplementary Notes		

16. Abstract

The "100-Car Naturalistic Driving Study" is a three-phased effort designed to accomplish three objectives: Phase I, Conduct Test Planning Activities; Phase II, Conduct a Field Test; and Phase III, Prepare for Large-Scale Field Data Collection Effort. This report documents the efforts of Phase II. Project sponsors are the National Highway Traffic Safety Administration (NHTSA) and the Virginia Department of Transportation (VDOT).

The 100-Car Naturalistic Driving Study is the first instrumented-vehicle study undertaken with the primary purpose of collecting large-scale, naturalistic driving data. Drivers were given no special instructions, no experimenter was present, and the data collection instrumentation was unobtrusive. In addition, 78 of 100 vehicles were privately owned. The resulting database contains many extreme cases of driving behavior and performance, including severe drowsiness, impairment, judgment error, risk taking, willingness to engage in secondary tasks, aggressive driving, and traffic violations. The data set includes approximately 2,000,000 vehicle miles, almost 43,000 hours of data, 241 primary and secondary drivers, 12 to 13 months of data collection for each vehicle, and data from a highly capable instrumentation system including 5 channels of video and many vehicle state and kinematic sensors. From the data, an "event" database was created, similar in classification structure to an epidemiological crash database, but with video and electronic driver and vehicle performance data. The events are crashes, near-crashes, and other "incidents." Data is classified by pre-event maneuver, precipitating factor, event type, contributing factors, associative factors, and the avoidance maneuver. Parameters such as vehicle speed, vehicle headway, time-to-collision, and driver reaction time are also recorded.

The current project specified ten objectives or *goals* that would be addressed through the initial analysis of the event database. This report addresses the first 9 of these goals, which include analyses of rear-end events, lane change events, the role of inattention, and the relationship between levels of severity. Goal 10 is a separate report and addresses the implications for a larger-scale data collection effort.

17. Key Words 100-Car, Naturalistic, Rear-End Collision, Lane Change, Collision Avoidance, Intelligent Vehicle Initiative, Driver Behavior, Human Factors, Inattention, Distraction, Lead-Vehicle Conflict		18. Distribution Statement This document is available to the public through the National Technical Information Service, Springfield, VA 22161	
19. Security Classif. (of this report) Unclassified	20. Security Classif. (of this page) Unclassified	21. No. of Pages 422	22. Price

Form DOT F 1700.7 (8-72)

This page intentionally left blank.

TABLE OF CONTENTS

GLOSSARY OF TERMS ..... **xviii**

EXECUTIVE SUMMARY ..... **xxii**

REPORT OVERVIEW ..... **xxvi**

INTRODUCTION xxvi
METHOD xxviii
Instrumentation ..... *xxviii*
Subjects ..... *xxxii*
Vehicles ..... *xxxv*
Results for the Initial 10 Project Goals ..... *xxxv*
GOAL 1: CHARACTERIZATION OF CRASHES, NEAR-CRASHES AND INCIDENTS FOR THE 100-CAR STUDY xxxv
GOAL 2: QUANTIFICATION OF NEAR-CRASH EVENTS xxxix
GOAL 3: CHARACTERIZATION OF DRIVER INATTENTION xl
GOAL 4: DRIVER BEHAVIOR OVER TIME xliii
Driving Behavior Over the Course of a Year ..... *xliv*
Driving Behavior Over the First 50 Hours ..... *xlv*
Performance for Same Driver for Four Weeks in Private and Leased Vehicles ..... *xlvi*
GOAL 5: REAR-END CONFLICT CAUSAL FACTORS AND DYNAMIC CONDITIONS xlvii
Following Vehicle Data ..... *xlviii*
GOAL 6: LANE CHANGE CAUSAL FACTORS AND DYNAMIC CONDITIONS xlix
GOAL 7: INATTENTION FOR REAR END LEAD-VEHICLE SCENARIOS lii
GOAL 8: CHARACTERIZE THE REAR END SCENARIOS IN RELATION TO HEINRICH'S TRIANGLE lv
GOAL 9: EVALUATE PERFORMANCE OF HARDWARE, SENSORS, AND DATA COLLECTION SYSTEM lvii

CHAPTER 1: INTRODUCTION **1**

BACKGROUND 1
PHASE I. CONDUCT TEST PLANNING ACTIVITIES 2
Task 1: Establish Intelligent Vehicle Initiative (IVI) Data Needs *2*
Task 2: Develop Phase I Test Requirements *4*
Task 3: Select Candidate Test Areas and Evaluate Crash Frequency Data *4*
Task 4: Determine Crash Sampling Requirements *4*
Task 5: Determine Driver/Vehicle Demographic Requirements *4*
Task 6: Determine Near-Crash Statistical Power Requirements *4*
Task 7: Conduct Trade Study – Research Design Parameters/ Sampling Rates/Formats Concept *5*
Task 8: Conduct Trade Study to Determine Vehicle Types *5*
Task 9: Develop Participant Recruiting Specification *5*
Task 10: Develop Test Data Collection Plan *6*
Task 11: Develop Test Reduction, Archiving, and Analysis Plan *6*
Task 12: Development of Data Collection System Requirements *7*
Task 13: Review/Test of Technology/Sensor Alternatives *7*
Task 14: Review/Test of Trigger Criteria Methods *7*
Task 15: Trade Study Analysis of Hardware/Software Alternatives *7*

CHAPTER 2: METHOD FOR PHASE II – THE 100-CAR FIELD TEST **9**

DRIVERS 9
THE 100-CAR DATA ACQUISTION SYSTEM 9
Node 1: Vehicle Network Box *11*
Node 2: Accelerometer Box *12*
Nodes 3-4: Headway Detection *12*
Node 5: Side Obstacle Detection *12*
Node 6: GPS Data Node *12*
Node 7: Automatic Collision Notification *13*
Node 8: Cellular Communications *13*
Node 9: Incident Box *13*
Node 10: System Initialization. *14*

Lane Tracking System *14*
Video Data *14*
Driving Performance Data *16*
DATA COLLECTION, ARCHIVAL AND STORAGE 17
Demographic and Questionnaire Data *17*
Instrumentation *19*
Data Retrieval and Storage *19*
PROCEDURE FOR DATA REDUCTION 19
Sensitivity Analysis *19*
Recruiting and Training Data Reductionists *22*
Data Reduction Software Framework *23*
Data Reduction Inter- and Intra-Rater Reliability *24*
Database Creation *26*
BY THE NUMBERS – TOP LEVEL PROJECT STATISTICS 26

CHAPTER 3: DRIVER AND VEHICLE DEMOGRAPHICS 31

AGE AND GENDER 31
Self-Reported Years of Driving Experience *32*
Driver Age as Compared to Self-Reported Years of Driving Experience *33*
Ethnicity *34*
Occupation *35*
Education *37*
Self-Reported Violations *38*
Self-Reported Crashes *41*
Self-Reported Annual Mileage *43*
Vehicle Demographics *44*

CHAPTER 4: LESSONS LEARNED IN CONDUCTING THE 100-CAR STUDY 47

BACKGROUND 47
SUBJECT RECRUITMENT AND COMPLIANCE 47
DAS INSTALLATION 48
HARDWARE AND SOFTWARE MAINTENANCE 49
Hardware *50*
Software *51*
DATA DOWNLOADING 52
DATA REDUCTION 53
DATA ANALYSIS 54
OTHER LOGISTICS 55
DISCUSSION 56

CHAPTER 5: GOAL 1, CLASSIFY AND QUANTIFY CAUSAL FACTORS AND DYNAMIC SCENARIOS INVOLVED IN EACH CONFLICT CATEGORY. 57

DATA ANALYSIS OVERVIEW 57

CHAPTER 6: GOAL 2, OPERATIONALLY DEFINE A NEAR-CRASH USING QUANTITATIVE MEASURES 139

DATA ANALYSIS OVERVIEW 139
Data Included in the Analyses *140*
DISCUSSION 155

CHAPTER 7: GOAL 3, CHARACTERIZATION OF DRIVER INATTENTION AS IT RELATES TO INCIDENTS, NEAR-CRASHES, AND CRASHES 157

DATA ANALYSIS OVERVIEW 157
Driver Data Included in the Analyses *158*
DISCUSSION 184

CHAPTER 8: GOAL 4, DRIVER PERFORMANCE IN INSTRUMENTED VEHICLES OVER TIME: OVER THE FIRST FEW HOURS, OVER THE FIRST YEAR, AND OVER ONE MONTH FOR SAME DRIVER IN PRIVATE VERSUS LEASED VEHICLES 187

DATA ANALYSIS OVERVIEW 187
Data Included in the Analyses 188
DISCUSSION 236
Question 1 and Question 4: Driving behavior over the year 237
Question 2 and Question 5: Driving behavior over the first 50 hours 237
Question 3 and Question 6: Same driver for four weeks in private and leased vehicles 238

CHAPTER 9: GOAL 5, DETERMINE REAR-END CRASH CONTRIBUTING FACTORS AND DYNAMIC CONDITIONS 240

DATA ANALYSIS OVERVIEW 240
Data Included in the Analyses 240
Following Vehicle Data 251
Lead-vehicle Data 254
Following Vehicle Data 257
DISCUSSION 263

CHAPTER 10: GOAL 6, DETERMINE LANE CHANGE CONTRIBUTING FACTORS AND DYNAMIC CONDITIONS 268

BACKGROUND 268
DATA ANALYSIS OVERVIEW 268
Data Included in the Analyses 268
DISCUSSION 280

CHAPTER 11: GOAL 7, DETERMINE THE DISTRIBUTION OF INATTENTION FOR EACH REAR-END LEAD-VEHICLE SCENARIO 284

DATA ANALYSIS OVERVIEW 284
Data Included in the Analyses 285
Lead-Vehicle Scenario Analysis 289
DISCUSSION 307

CHAPTER 12: GOAL 8, CHARACTERIZE EACH OF THE 4 RE SCENARIOS IN RELATION TO HEINRICH'S TRIANGLE 308

BACKGROUND 308
DATA ANALYSIS OVERVIEW 308
Data Included in the Analysis 309
DISCUSSION 314

CHAPTER 13: GOAL 9, EVALUATE THE PERFORMANCE OF THE HARDWARE, SENSORS, AND DATA COLLECTION SYSTEM USED IN PHASE II 316

DATA ANALYSIS OVERVIEW 316
Data Included in the Analyses 316
DISCUSSION 345

CHAPTER 14: GENERAL CONCLUSIONS 347

ADVANTAGES TO THE NATURALISTIC APPROACH 347
IMPLICATIONS FOR A LARGE-SCALE EFFORT 348
SUMMARY 349

REFERENCES 351

This page intentionally left blank.

LIST OF FIGURES

FIGURE RO.1. THE RELATIONSHIP BETWEEN EMPIRICAL, NATURALISTIC, AND EPIDEMIOLOGICAL METHODS IN DRIVING SAFETY RESEARCH. XXVII
FIGURE RO.2. A COMPRESSED VIDEO IMAGE FROM THE 100-CAR STUDY DATA. THE DRIVER'S FACE (UPPER LEFT QUADRANT) HAS BEEN DISTORTED TO PROTECT THE DRIVER'S IDENTITY. THE LOWER RIGHT QUADRANT IS SPLIT BETWEEN THE LEFT-SIDE VIEW (TOP) AND THE REAR VIEW (BOTTOM). XXIX
FIGURE RO.3. THE MAIN DATA ACQUISITION SYSTEM (DAS) UNIT MOUNTED UNDER THE "PACKAGE SHELF" OF THE TRUNK. XXX
FIGURE RO.4. THE 100-CAR STUDY DAS MAIN UNIT SHOWN WITHOUT THE TOP AND FRONT COVERS. XXX
FIGURE RO.5. DOPPLER RADAR ANTENNA MOUNTED ON THE FRONT OF A VEHICLE, COVERED BY ONE OF THE PLASTIC LICENSE PLATES USED FOR THIS STUDY. XXXI
FIGURE RO.6. THE INCIDENT PUSH BUTTON BOX MOUNTED ABOVE THE REARVIEW MIRROR. THE PORTION ON THE RIGHT CONTAINS THE DRIVER FACE/LEFT ROAD VIEW CAMERA HIDDEN BY A SMOKED PLEXIGLAS COVER. XXXI
FIGURE RO.7. THE MOUNTING FOR THE GLARE SENSOR BEHIND THE REARVIEW MIRROR. NOTE THE FORWARD VIEW CAMERA AS PART OF THE SAME MOUNTING ASSEMBLY. XXXII
FIGURE RO.8. NUMBER OF SELF-REPORTED TRAFFIC VIOLATIONS IN THE PAST 5 YEARS AS A PERCENTAGE OF DRIVER AGE GROUP. XXXIV
FIGURE RO.9. NUMBER OF SELF-REPORTED TRAFFIC CRASHES IN THE PAST 10 YEARS AS A PERCENTAGE OF DRIVER AGE GROUP. XXXIV
FIGURE RO.10. DISTRIBUTION OF VEHICLE MAKES AND MODELS DRIVEN DURING THE STUDY. XXXV
FIGURE RO.12. POINT OF GREATEST THREAT WITH LEAD VEHICLE FOR ALL CRASHES AND NEAR-CRASHES, AND A RANDOM SAMPLE OF INVALID EVENTS. THE BOUNDARIES SHOWN ARE APPROXIMATIONS OF WARNING AND CONFLICT BOUNDARIES USED AS PART OF A FORWARD COLLISION WARNING ALGORITHM (KIEFER ET AL., 2003) AND A MINIMUM ERROR BOUNDARY CALCULATED FOR THIS DATASET. XL
FIGURE RO.13. PERCENTAGE OF EVENTS FOR ATTENTION BY SEVERITY LEVEL. XLI
FIGURE RO.14. COMPARISON OF CRASHES AND NEAR-CRASHES THE FREQUENCY OF OCCURRENCES OF THE PRESENCE OF A DISTRACTING AGENT AS A CONTRIBUTING FACTOR. XLII
FIGURE RO.15. FREQUENCY OF EVENTS FOR WHICH THE CONTRIBUTING FACTOR WAS WIRELESS DEVICE USE. XLIII
FIGURE RO.16. MEAN NUMBER OF EVENTS FOR A MATCHED SET OF YOUNGER DRIVERS FOR LEASED AND PRIVATE VEHICLES OVER WEEKS 1-50 OF THE STUDY. XLV
FIGURE RO.17. MATCHED SET OF YOUNGER SWITCH DRIVERS: LEASED VERSUS PRIVATE VEHICLE MEAN NUMBER OF EVENTS FOR WEEKS 1-4. XLVII
FIGURE RO.18. RATE PER MVMT FOR SV STRIKING BY AGE AND LANE CHANGE MANEUVER FOR NEAR-CRASHES. LI
FIGURE RO.19. RATE PER MVMT FOR STRUCK SV BY AGE AND LANE CHANGE MANEUVER FOR INCIDENTS. LI
FIGURE RO.20. PERCENTAGE OF LEAD-VEHICLE EVENTS FOR WHICH INATTENTION WAS LISTED AS A CONTRIBUTING FACTOR (INCLUDES THE NONSPECIFIC EYEGLANCE EVENTS FOR CRASHES AND NEAR-CRASHES). LIII
FIGURE RO.21. FREQUENCY OF CRASHES BY DRIVER ATTENTION AND LEAD-VEHICLE KINEMATIC SCENARIO. LIV
FIGURE RO.22. FREQUENCY OF NEAR-CRASHES BY DRIVER ATTENTION LEVEL AND LV SCENARIO. LIV
FIGURE RO.23. FREQUENCY OF SECONDARY TASK INATTENTION SOURCES FOR LEAD-VEHICLE EVENTS. LV
FIGURE RO.24. HEINRICH'S TRIANGLE FOR REAR-END CONFLICTS IN THE 100-CAR STUDY. LVI
FIGURE 2.1. 100-CAR STUDY DATA HARDWARE COLLECTION SYSTEM SCHEMATIC. 10
FIGURE 2.2. THE MAIN DAS UNIT MOUNTED UNDER THE "PACKAGE SHELF" OF THE TRUNK. 11
FIGURE 2.3. THE 100-CAR STUDY DAS. 11
FIGURE 2.4. COMPUTER BOARD FOR THE VORAD UNIT. 12
FIGURE 2.5. RADAR UNIT MOUNTED ON THE FRONT OF A VEHICLE, COVERED BY A LICENSE PLATE. 12
FIGURE 2.6. THE INCIDENT PUSHBUTTON BOX MOUNTED ABOVE THE REARVIEW MIRROR. 13
FIGURE 2.7. THE MOUNTING FOR THE GLARE SENSOR BEHIND THE REARVIEW MIRROR. NOTE THE FORWARD VIEW CAMERA AS PART OF THE SAME MOUNTING ASSEMBLY. 14
FIGURE 2.8. THE 5 CAMERA VIEWS RECORDED IN THE INSTRUMENTED VEHICLE: (1) FORWARD, (2) DRIVER'S FACE/LEFT SIDE OF VEHICLE, (3) REAR-FACING, (4) OVER THE DRIVER'S SHOULDER CAPTURING THE DRIVER'S HANDS AND FEET, THE STEERING WHEEL, AND THE INSTRUMENT PANEL, AND (5) RIGHT SIDE OF VEHICLE. 15
FIGURE 2.9. THE DOUBLE QUAD, SPLIT VIDEO IMAGE. 15
FIGURE 2.10. A VIDEO IMAGE FROM THE 100-CAR STUDY DATA. THE DRIVER'S FACE HAS BEEN DISTORTED TO PROTECT THE DRIVER'S IDENTITY. 16
FIGURE 2.11. FLOW CHART OF THE DATA REDUCTION PROCESS. 20

FIGURE 2.12. GRAPHICAL DEPICTION OF TRIGGER CRITERIA SETTINGS FOR PHASE II AND PHASE IV USING THE DISTRIBUTION OF VALID EVENTS. NOTE THAT THIS DISTRIBUTION AND CRITERION PLACEMENT IS UNIQUE FOR EACH TRIGGER TYPE. 21
FIGURE 2.13. DECISION TREE USED TO CLASSIFY EVENT SEVERITY. 23
FIGURE 3.1. DISTRIBUTION OF PARTICIPANT AGE. 32
FIGURE 3.2. DISTRIBUTION OF PARTICIPANT SELF-REPORTED YEARS OF DRIVING EXPERIENCE. 33
FIGURE 3.3. DISTRIBUTION OF PARTICIPANTS ACROSS SELF-REPORTED YEARS OF DRIVING EXPERIENCE BINS. 34
FIGURE 3.4. DISTRIBUTION OF PARTICIPANTS ACROSS SELF-REPORTED ETHNIC GROUPS. 35
FIGURE 3.5. DISTRIBUTION OF PARTICIPANTS ACROSS JOB CATEGORIES. 37
FIGURE 3.6. DISTRIBUTION OF PARTICIPANTS WITH REGARD TO YEARS AND TYPE OF EDUCATION. 38
FIGURE 3.7. DISTRIBUTION OF TYPE OF SELF-REPORTED TRAFFIC VIOLATIONS IN THE PAST 5 YEARS. 39
FIGURE 3.8. PERCENTAGE OF PARTICIPANTS BY GENDER FOR NUMBER OF SELF-REPORTED TRAFFIC VIOLATIONS IN THE PAST 5 YEARS. 40
FIGURE 3.9. PERCENTAGE OF PARTICIPANTS IN EACH AGE BIN FOR NUMBER OF SELF-REPORTED TRAFFIC VIOLATIONS IN THE PAST 5 YEARS. 41
FIGURE 3.10. PERCENTAGE OF PARTICIPANTS BY GENDER FOR NUMBER OF SELF-REPORTED TRAFFIC CRASHES IN THE PAST 10 YEARS. 42
FIGURE 3.11. PERCENTAGE OF PARTICIPANTS IN EACH AGE BIN FOR NUMBER OF SELF-REPORTED TRAFFIC CRASHES IN THE PAST 10 YEARS. 42
FIGURE 3.12. DISTRIBUTION OF PARTICIPANTS ACROSS SELF-REPORTED ANNUAL MILEAGE CATEGORIES. 43
FIGURE 3.13. PERCENTAGE OF PARTICIPANTS FOR EACH AGE GROUP DRIVING PERSONAL VERSUS LEASED VEHICLES FOR THE STUDY. 44
FIGURE 3.14. DISTRIBUTION OF PARTICIPANTS ACROSS VEHICLE MAKES/MODELS DRIVEN DURING THE STUDY. 45
FIGURE 5.1. EXAMPLE OF THE RELATIONSHIP BETWEEN THE ANALYZED FACTORS FOR THE 100-CAR STUDY. 58
FIGURE 5.2. PRE-EVENT MANEUVERS FOR SINGLE VEHICLE CRASHES 64
FIGURE 5.3. BREAKDOWN OF DRIVING ENVIRONMENT VARIABLES FOR SINGLE VEHICLE CRASHES. 65
FIGURE 5.4. BREAKDOWN OF PRECIPITATING FACTORS FOR NEAR-CRASHES INVOLVING A SINGLE VEHICLE. 66
FIGURE 5.5. BREAKDOWN OF SECONDARY TASK DISTRACTIONS FOR NEAR-CRASHES INVOLVING A SINGLE VEHICLE. 67
FIGURE 5.6. BREAKDOWN OF DRIVING ENVIRONMENT FOR INCIDENTS INVOLVING SINGLE VEHICLES. 69
FIGURE 5.7. BREAKDOWN OF INFRASTRUCTURE VARIABLES FOR INCIDENTS INVOLVING SINGLE VEHICLES. 70
FIGURE 5.8. THE PRECIPITATING FACTORS FOR INCIDENTS INVOLVING SINGLE VEHICLES. 71
FIGURE 5.9. BREAKDOWN OF SECONDARY TASKS FOR INCIDENTS INVOLVING SINGLE VEHICLES. 72
FIGURE 5.10. BREAKDOWN OF THE DRIVING ENVIRONMENT VARIABLES FOR INCIDENTS INVOLVING SINGLE VEHICLES. 73
FIGURE 5.11. BREAKDOWN OF THE INFRASTRUCTURE VARIABLES FOR INCIDENTS INVOLVING SINGLE VEHICLES. 74
FIGURE 5.12. BREAKDOWN OF THE SECONDARY TASKS CONTRIBUTING TO CRASHES INVOLVING A LEAD VEHICLE. 76
FIGURE 5.13. BREAKDOWN OF THE DRIVING ENVIRONMENT VARIABLES FOR CRASHES WITH LEAD VEHICLES. 77
FIGURE 5.14. BREAKDOWN OF INFRASTRUCTURE VARIABLES FOR CRASHES WITH LEAD VEHICLES. 78
FIGURE 5.15. BREAKDOWN OF PRECIPITATING FACTOR FOR NEAR-CRASHES INVOLVING LEAD VEHICLES. 81
FIGURE 5.16. BREAKDOWN OF SECONDARY TASKS FOR NEAR-CRASHES INVOLVING A LEAD VEHICLE. 82
FIGURE 5.19. BREAKDOWN OF ALL PRE-EVENT MANEUVERS THAT OCCURRED PRIOR TO INCIDENTS INVOLVING LEAD VEHICLES. 86
FIGURE 5.20. BREAKDOWN OF PRECIPITATING FACTORS FOR INCIDENTS INVOLVING LEAD VEHICLES. 87
FIGURE 5.21. BREAKDOWN OF INATTENTION CATEGORIES FOR INCIDENTS INVOLVING A LEAD VEHICLE. 88
FIGURE 5.22. BREAKDOWN OF WIRELESS DEVICE OPERATIONS THAT CONTRIBUTED TO INCIDENTS INVOLVING LEAD VEHICLES. 89
FIGURE 5.23. BREAKDOWN OF DRIVER FACTORS THAT CONTRIBUTE TO INCIDENTS INVOLVING LEAD VEHICLES. 90
FIGURE 5.24. BREAKDOWN OF AVOIDANCE MANEUVERS THAT OCCURRED DURING INCIDENTS INVOLVING LEAD VEHICLES. 91
FIGURE 5.25. BREAKDOWN OF THE DRIVING ENVIRONMENT VARIABLES THAT CONTRIBUTE TO INCIDENTS INVOLVING LEAD VEHICLES. 92
FIGURE 5.26. BREAKDOWN OF DRIVING ENVIRONMENT VARIABLES FOR CRASHES INVOLVING FOLLOWING VEHICLES. 95
FIGURE 5.27. BREAKDOWN FOR PRECIPITATING FACTORS FOR NEAR-CRASHES INVOLVING FOLLOWING VEHICLES. 96
FIGURE 5.28. BREAKDOWN OF DRIVING ENVIRONMENT VARIABLES FOR NEAR-CRASHES INVOLVING FOLLOWING VEHICLES. 97

FIGURE 5.29. BREAKDOWN OF INFRASTRUCTURE-RELATED VARIABLES FOR NEAR-CRASHES INVOLVING FOLLOWING VEHICLES. 98
FIGURE 5.30. BREAKDOWN FOR PRECIPITATING FACTORS FOR INCIDENTS INVOLVING FOLLOWING VEHICLES. 100
FIGURE 5.32. BREAKDOWN OF INFRASTRUCTURE VARIABLES FOR INCIDENTS INVOLVING FOLLOWING VEHICLES. ...103
FIGURE 5.33. BREAKDOWN OF DRIVING ENVIRONMENT VARIABLES INVOLVING CRASHES WITH OBSTACLES/OBJECTS. 105
FIGURE 5.34. BREAKDOWN OF INFRASTRUCTURE-RELATED VARIABLES FOR CRASHES INVOLVING OBJECTS OR OBSTACLES. 106
FIGURE 5.35. BREAKDOWN OF SECONDARY TASKS FOR INCIDENTS INVOLVING OBSTACLES OR OBJECTS. 109
FIGURE 5.36. BREAKDOWN OF DRIVING ENVIRONMENT VARIABLES FOR INCIDENTS INVOLVING OBJECTS OR OBSTACLES. 110
FIGURE 5.37. BREAKDOWN OF INFRASTRUCTURE-RELATED VARIABLES FOR INCIDENTS INVOLVING OBSTACLES OR OBJECTS. 111
FIGURE 5.38. BREAKDOWN OF PRECIPITATING FACTORS FOR INCIDENTS INVOLVING A PARKED VEHICLE. 115
FIGURE 5.39. BREAKDOWN OF SECONDARY TASKS FOR INCIDENTS INVOLVING A PARKED VEHICLE. 116
FIGURE 5.40. BREAKDOWN OF SECONDARY TASKS FOR NEAR-CRASHES INVOLVING A VEHICLE IN THE ADJACENT LANE. 120
FIGURE 5.41. BREAKDOWN OF SECONDARY TASKS FOR INCIDENTS INVOLVING A CONFLICT WITH VEHICLE IN THE ADJACENT LANE. 123
FIGURE 6.1. POINT OF GREATEST THREAT WITH LEAD VEHICLE FOR ALL CRASHES AND NEAR-CRASHES, AND A RANDOM SAMPLE OF INVALID EVENTS. THE BOUNDARIES SHOWN ARE APPROXIMATIONS OF THE WARNING AND CONFLICT BOUNDARIES USED AS PART OF THE FORWARD COLLISION WARNING ALGORITHM (KIEFER ET AL., 2003) AND A MINIMUM ERROR BOUNDARY CALCULATED FOR THIS DATASET. 143
FIGURE 6.2. RANGE/RANGE RATE TRAJECTORIES OF VEHICLES APPROACHING A LEAD VEHICLE INCLUDING CRASHES AND NEAR-CRASHES, AND A RANDOM SAMPLE OF INVALID EVENTS. THE BOUNDARIES SHOWN ARE APPROXIMATIONS OF THE WARNING AND CONFLICT BOUNDARIES USED AS PART OF THE FORWARD COLLISION WARNING ALGORITHM (KIEFER ET AL., 2003) AND A MINIMUM ERROR BOUNDARY CALCULATED FOR THIS DATASET. 144
FIGURE 6.3. RANGE/RANGE RATE TRAJECTORIES OF VEHICLES APPROACHING A LEAD VEHICLE. THIS DATA INCLUDES A RANDOM SAMPLE OF CRASHES AND NEAR-CRASHES, AND INVALID EVENTS TO IMPROVE VISUALIZATION. THE BOUNDARIES SHOWN ARE APPROXIMATIONS OF THE WARNING AND CONFLICT BOUNDARIES USED AS PART OF THE FORWARD COLLISION WARNING ALGORITHM (KIEFER ET AL., 2003) AND A MINIMUM ERROR BOUNDARY CALCULATED FOR THIS DATASET. 145
FIGURE 6.4. RANGE/RANGE RATE TRAJECTORIES FOR ALL CRASHES. TRAJECTORIES INCLUDE THE DATA POINTS UP TO 2 SECONDS PRIOR TO THE CRASH. THE BOUNDARIES SHOWN ARE APPROXIMATIONS OF THE WARNING AND CONFLICT BOUNDARIES USED AS PART OF THE FORWARD COLLISION WARNING ALGORITHM (KIEFER ET AL., 2003) AND A MINIMUM ERROR BOUNDARY CALCULATED FOR THIS DATASET. 146
FIGURE 6.5. POINT OF GREATEST THREAT WITH FOLLOWING VEHICLE FOR ALL CRASHES AND NEAR-CRASHES, AND A RANDOM SAMPLE OF INVALID EVENTS. THE BOUNDARY SHOWN IS THE MINIMUM ERROR BOUNDARY CALCULATED FOR THIS DATASET. 148
FIGURE 6.6. RANGE/RANGE RATE TRAJECTORIES FOR FOLLOWING VEHICLES INCLUDING CRASHES AND NEAR-CRASHES, AND A RANDOM SAMPLE OF INVALID EVENTS. THE BOUNDARY SHOWN IS A MINIMUM ERROR BOUNDARY CALCULATED FOR THIS DATASET. 149
FIGURE 6.7. RANGE/RANGE RATE TRAJECTORIES FOR FOLLOWING VEHICLES INCLUDING A RANDOM SAMPLE OF CRASHES AND NEAR-CRASHES, AND INVALID EVENTS. THE BOUNDARY SHOWN IS THE MINIMUM ERROR BOUNDARY CALCULATED FOR THIS DATASET. 150
FIGURE 6.8. RANGE/RANGE RATE TRAJECTORIES FOR ALL FOLLOWING-VEHICLE CRASH EVENTS. THE CRASH PREDICTION TRAJECTORY SHOWN INCLUDES THE DATA POINTS UP TO 2 SECONDS PRIOR TO THE CRASH. THE BOUNDARY SHOWN IS A MINIMUM ERROR BOUNDARY CALCULATED FOR THIS DATASET. 151
FIGURE 7.1. NUMBER OF EVENTS PER MVMT (N=98). 159
FIGURE 7.2. FREQUENCY OF INATTENTION-RELATED EVENTS PER MVMT IN WHICH INATTENTION IS DUE TO: (1) DROWSINESS, (2) INATTENTION TO THE FORWARD ROADWAY, (3) SECONDARY TASK, (4) SPECIFIC EYEGLANCE AWAY FROM FORWARD ROADWAY (FOR CRASHES AND NEAR-CRASHES ONLY) OR (5) NONSPECIFIC EYEGLANCE AWAY FROM FORWARD ROADWAY (FOR CRASHES AND NEAR-CRASHES ONLY). 160
FIGURE 7.3. FREQUENCY OF CRASHES, NEAR-CRASHES, AND INCIDENTS FOR WHICH DRIVERS WERE INATTENTIVE VERSUS ATTENTIVE (DRIVER N = 241). 161

FIGURE 7.4. PERCENTAGE OF EVENTS THAT DRIVERS WERE INATTENTIVE VERSUS ATTENTIVE BY SEVERITY LEVEL (DRIVER N = 241). 162
FIGURE 7.5. THE FREQUENCY OF CRASHES AND NEAR-CRASHES IN WHICH THESE TYPES OF INATTENTION WERE IDENTIFIED AS A CONTRIBUTING FACTOR. 164
FIGURE 7.6. THE PERCENTAGE OF CRASHES AND NEAR-CRASHES IN WHICH THESE TYPES OF INATTENTION WERE IDENTIFIED AS A CONTRIBUTING FACTOR. 164
FIGURE 7.7. THE FREQUENCY OF INCIDENTS IN WHICH THESE TYPES OF INATTENTION WERE IDENTIFIED AS A CONTRIBUTING FACTOR. 165
FIGURE 7.8. COMPARISON OF INATTENTION AND ATTENTION-RELATED EVENTS BY SEVERITY LEVEL PER MVMT (DRIVER N = 98). 166
FIGURE 7.9. THE RATE OF INATTENTIVE VERSUS ATTENTIVE EVENTS PER MVMT BY AGE GROUP. 167
FIGURE 7.10. THE RATE OF INATTENTIVE VERSUS ATTENTIVE CRASHES PER MVMT BY AGE GROUP. 168
FIGURE 7.11. THE RATE OF INATTENTIVE VERSUS ATTENTIVE NEAR-CRASHES PER MVMT BY AGE GROUP. 168
FIGURE 7.12. COMPARISON OF THE NUMBER OF OCCURRENCES OF THE PRESENCE DISTRACTING AGENT AS A CONTRIBUTING FACTOR (DRIVER N = 241). 169
FIGURE 7.13. COMPARISON OF CRASHES AND NEAR-CRASHES THE FREQUENCY OF OCCURRENCES OF THE PRESENCE, EITHER ALONE OR IN COMBINATION, OF THE DISTRACTING AGENT AS A CONTRIBUTING FACTOR (DRIVER N = 241). 170
FIGURE 7.14. FREQUENCY OF OCCURRENCE OF DRIVING-RELATED INATTENTION TO FORWARD ROADWAY, ALONE OR IN COMBINATION, BY LEVEL OF SEVERITY (DRIVER N = 241). 171
FIGURE 7.15. FREQUENCY OF OCCURRENCES IN WHICH THE CONTRIBUTING FACTOR WAS WIRELESS DEVICE USE (DRIVER N = 241). 172
FIGURE 7.16. FREQUENCY OF OCCURRENCES FOR WHICH THE CONTRIBUTING FACTOR WAS A VEHICLE-RELATED TASK. 173
FIGURE 7.17. FREQUENCY OF OCCURRENCES IN WHICH THE CONTRIBUTING FACTOR WAS PASSENGER-RELATED SECONDARY TASKS. (DRIVER N = 241). 174
FIGURE 7.18. FREQUENCY OF OCCURRENCES IN WHICH THE CONTRIBUTING FACTOR WAS TALKING/SINGING: NO PASSENGER APPARENT SECONDARY TASKS. (DRIVER N = 241). 175
FIGURE 7.19. FREQUENCY OF OCCURRENCES IN WHICH THE CONTRIBUTING FACTOR WAS AN EXTERNAL SECONDARY TASK (DRIVER N = 241). 176
FIGURE 7.20. THE FREQUENCY OF OCCURRENCES IN WHICH THE CONTRIBUTING FACTOR WAS INTERNAL SECONDARY TASKS (DRIVER N = 241). 177
FIGURE 7.21. FREQUENCY OF OCCURRENCES IN WHICH THE CONTRIBUTING FACTOR WAS A PERSONAL HYGIENE TASK (DRIVER N = 241). 178
FIGURE 7.22. FREQUENCY OF OCCURRENCES IN WHICH THE CONTRIBUTING FACTOR WAS A DINING TASK (DRIVER N = 241). 179
FIGURE 7.23. FREQUENCY OF OCCURRENCES IN WHICH THE CONTRIBUTING FACTOR WAS A SMOKING TASK (DRIVER N = 241). 180
FIGURE 7.24. FREQUENCY OF OCCURRENCES IN WHICH THE CONTRIBUTING FACTOR WAS A DAYDREAMING TASK (DRIVER N = 241). 181
FIGURE 7.25. FREQUENCY OF DRIVER INATTENTION, WILLFUL BEHAVIOR, AND DRIVER IMPAIRMENT ON THE TOTAL NUMBER OF CRASHES, NEAR-CRASHES, AND INCIDENTS IN THE 100-CAR STUDY DATABASE. 182
FIGURE 7.26. FREQUENCY OF THE DRIVER PHYSICAL/MENTAL IMPAIRMENT CATEGORIES LISTED AS A CONTRIBUTING FACTOR. 183
FIGURE 7.27. THE PERCENTAGE OF EVENTS THAT INCLUDED BOTH INATTENTION AND WILLFUL BEHAVIOR, DRIVER PROFICIENCY, OR DRIVER IMPAIRMENT. 184
FIGURE 8.1. MEAN NUMBER OF EVENTS PER VEHICLE FOR WEEKS 1-50 FOR LEASED AND PRIVATE VEHICLES. 190
FIGURE 8.2. COMPARISON OF RR USING THE ORIGINALLY PROPOSED COMPARISON WEEKS OF 1, 4, 12, 26, AND 49 USING WEEK 50 AS A CONTROL WEEK. 190
FIGURE 8.3. COMPARISON OF RR USING WEEKS OF 3, 7, 18, 27, AND 49 USING WEEK 50 AS A CONTROL WEEK AND SHOWING A STRONG DOWNWARD TREND. 191
FIGURE 8.4. COMPARISON OF RR USING WEEKS OF 1, 5, 13, 33, AND 49 USING WEEK 50 AS A CONTROL TIME PERIOD AND SHOWING AN UPWARD TREND FOLLOWED BY A DOWNWARD TREND. 192
FIGURE 8.5. MEAN NUMBER OF EVENTS PER VEHICLE FOR LEASED AND PRIVATE VEHICLES FOR TIME PERIODS OF INTEREST OVER WEEKS 1-50 OF THE STUDY. 194
FIGURE 8.6. RELATIVE RISK FOR LEASED VERSUS PRIVATE VEHICLES FOR TIME PERIODS OF INTEREST OVER WEEKS 1-50 OF THE STUDY. 195

FIGURE 8.7. RELATIVE RISK AND 95TH PERCENTAGE CONFIDENCE INTERVALS FOR LEASED VERSUS PRIVATE VEHICLES FOR TIME PERIODS OF INTEREST OVER WEEKS 1-50 OF THE STUDY....195
FIGURE 8.8. MEAN NUMBER OF EVENTS PER VEHICLE FOR YOUNGER AND OLDER DRIVERS FOR TIME PERIODS OF INTEREST OVER WEEKS 1-50 OF THE STUDY....197
FIGURE 8.9. RELATIVE RISK FOR YOUNGER VERSUS OLDER DRIVERS FOR TIME PERIODS OF INTEREST OVER WEEKS 1-50 OF THE STUDY....197
FIGURE 8.10. RELATIVE RISK AND 95TH PERCENTILE CONFIDENCE INTERVALS FOR LEASED VERSUS PRIVATE VEHICLES FOR TIME PERIODS OF INTEREST OVER WEEKS 1-50 OF THE STUDY....198
FIGURE 8.11. MEAN NUMBER OF EVENTS FOR YOUNGER DRIVERS FOR LEASED AND PRIVATE VEHICLES FOR TIME PERIODS OF INTEREST OVER WEEKS 1-50 OF THE STUDY....202
FIGURE 8.12. RELATIVE RISK FOR YOUNGER DRIVER LEASED VERSUS PRIVATE VEHICLES FOR TIME PERIODS OF INTEREST OVER WEEKS 1-50 OF THE STUDY....202
FIGURE 8.13. RELATIVE RISK AND 95TH PERCENTAGE CONFIDENCE INTERVALS FOR YOUNGER DRIVER LEASED VERSUS PRIVATE VEHICLES FOR TIME PERIODS OF INTEREST OVER WEEKS 1-50 OF THE STUDY....203
FIGURE 8.14. MEAN NUMBER OF NEAR-CRASHES FOR YOUNGER DRIVERS FOR LEASED AND PRIVATE VEHICLES FOR TIME PERIODS OF INTEREST OVER WEEKS 1-50 OF THE STUDY....205
FIGURE 8.15. RELATIVE RISK FOR NEAR-CRASHES FOR YOUNGER DRIVERS LEASED VERSUS PRIVATE VEHICLES FOR TIME PERIODS OF INTEREST OVER WEEKS 1-50 OF THE STUDY....205
FIGURE 8.16. RELATIVE RISK AND 95TH PERCENTAGE CONFIDENCE INTERVALS FOR NEAR-CRASHES FOR YOUNGER DRIVERS LEASED VERSUS PRIVATE VEHICLES FOR TIME PERIODS OF INTEREST OVER WEEKS 1-50....206
FIGURE 8.17. MEAN NUMBER OF CRASHES FOR YOUNGER DRIVERS FOR LEASED AND PRIVATE VEHICLES FOR TIME PERIODS OF INTEREST OVER WEEKS 1-50 OF THE STUDY....208
FIGURE 8.18. RELATIVE RISK FOR CRASHES FOR YOUNGER DRIVER LEASED VERSUS PRIVATE VEHICLES FOR TIME PERIODS OF INTEREST OVER WEEKS 1-50 OF THE STUDY....208
FIGURE 8.19. OVERALL TREND OF EVENTS OVER THE FIRST 50 HOURS OF DRIVING A NEWLY INSTRUMENTED VEHICLE....210
FIGURE 8.20. MEAN NUMBER OF EVENTS PER VEHICLE FOR LEASED AND PRIVATE VEHICLES FOR TIME PERIODS OF INTEREST OVER HOURS 1-50 OF THE STUDY....212
FIGURE 8.21. LEASED VERSUS PRIVATE VEHICLE RELATIVE RISK FOR TIME PERIODS OF INTEREST OVER HOURS 1-50 OF THE STUDY....212
FIGURE 8.22. MEAN NUMBER OF EVENTS PER VEHICLE FOR YOUNGER DRIVERS FOR LEASED AND PRIVATE VEHICLES FOR TIME PERIODS OF INTEREST OVER HOURS 1-50 OF THE STUDY....214
FIGURE 8.23. YOUNGER DRIVER LEASED VERSUS PRIVATE VEHICLE RELATIVE RISK FOR TIME PERIODS OF INTEREST OVER HOURS 1-50 OF THE STUDY....214
FIGURE 8.24. MEAN NUMBER OF NEAR-CRASHES PER VEHICLE FOR YOUNGER DRIVERS FOR LEASED AND PRIVATE VEHICLES FOR TIME PERIODS OF INTEREST OVER HOURS 1-50 OF THE STUDY....216
FIGURE 8.25. YOUNGER DRIVER LEASED VERSUS PRIVATE VEHICLE NEAR-CRASH RELATIVE RISK FOR TIME PERIODS OF INTEREST OVER HOURS 1-50 OF THE STUDY....216
FIGURE 8.26. MATCHED SET OF SWITCH DRIVERS LEASED VERSUS PRIVATE VEHICLE MEAN NUMBER OF EVENTS FOR WEEKS 1-4....218
FIGURE 8.27. MATCHED SET OF SWITCH DRIVERS LEASED VERSUS PRIVATE VEHICLE RELATIVE RISK FOR WEEKS 1-4....218
FIGURE 8.28. MATCHED SET OF SWITCH DRIVERS LEASED VERSUS PRIVATE VEHICLE RELATIVE RISK FOR WEEKS 1-4 SHOWN WITH 95TH PERCENTAGE UPPER AND LOWER CONFIDENCE INTERVALS....219
FIGURE 8.29. MATCHED SET OF YOUNGER SWITCH DRIVERS: LEASED VERSUS PRIVATE VEHICLE MEAN NUMBER OF EVENTS FOR WEEKS 1-4....220
FIGURE 8.30. MATCHED SET OF YOUNGER SWITCH DRIVERS: LEASED VERSUS PRIVATE VEHICLE RELATIVE RISK FOR WEEKS 1 -4....220
FIGURE 8.31. MATCHED SET OF YOUNGER SWITCH DRIVERS: LEASED VERSUS PRIVATE VEHICLE RELATIVE RISK FOR WEEKS 1-4 SHOWN WITH 95TH PERCENTAGE UPPER AND LOWER CONFIDENCE INTERVALS....221
FIGURE 8.32. MATCHED SET OF OLDER SWITCH DRIVERS: LEASED VERSUS PRIVATE VEHICLE MEAN NUMBER OF EVENTS FOR WEEKS 1-4....222
FIGURE 8.33. MATCHED SET OF OLDER SWITCH DRIVERS: LEASED VERSUS PRIVATE VEHICLE RELATIVE RISK FOR WEEKS 1-4....223
FIGURE 8.34. MATCHED SET OF OLDER SWITCH DRIVERS: LEASED VERSUS PRIVATE VEHICLE RELATIVE RISK FOR WEEKS 1-4 SHOWN WITH 95TH PERCENTAGE UPPER AND LOWER CONFIDENCE INTERVALS....224

FIGURE 8.35. MEAN NUMBER OF EVENTS PER MILE FOR LEASED AND PRIVATE VEHICLES FOR TIME PERIODS OF INTEREST OVER WEEKS 1-50 OF THE STUDY. 226
FIGURE 8.36. EVENTS PER MILE RELATIVE RISK FOR LEASED VERSUS PRIVATE VEHICLES FOR TIME PERIODS OF INTEREST OVER WEEKS 1-50 OF THE STUDY. 226
FIGURE 8.37. MEAN NUMBER OF EVENTS PER MILE FOR YOUNGER DRIVERS FOR LEASED AND PRIVATE VEHICLES FOR TIME PERIODS OF INTEREST OVER WEEKS 1-50 OF THE STUDY. 228
FIGURE 8.38. RELATIVE RISK FOR YOUNGER DRIVER LEASED VERSUS PRIVATE VEHICLES FOR TIME PERIODS OF INTEREST OVER WEEKS 1-50 OF THE STUDY. 228
FIGURE 8.39. MEAN NUMBER OF EVENTS PER MILE FOR LEASED AND PRIVATE VEHICLES FOR TIME PERIODS OF INTEREST OVER HOURS 1-50 OF THE STUDY. 230
FIGURE 8.40. EVENTS PER MILE LEASED VERSUS PRIVATE VEHICLE RELATIVE RISK FOR TIME PERIODS OF INTEREST OVER HOURS 1-50 OF THE STUDY. 231
FIGURE 8.41. MATCHED SET OF SWITCH DRIVERS LEASED VERSUS PRIVATE VEHICLE MEAN EVENTS PER MILE FOR WEEKS 1-4. 232
FIGURE 8.42. MATCHED SET OF SWITCH DRIVERS LEASED VERSUS PRIVATE VEHICLE RELATIVE RISK PER MILE FOR WEEKS 1-4. 233
FIGURE 8.43. NUMBER OF EVENTS FOR LEASED AND PRIVATE VEHICLE DRIVING IN THE FIRST 10 HOURS OF DRIVING FOR A MATCHED SET OF 6 SWITCH DRIVERS. 234
FIGURE 8.44. NUMBER OF EVENTS FOR LEASED AND PRIVATE VEHICLE DRIVING IN THE FIRST 10 HOURS OF DRIVING FOR A MATCHED SET OF 4 YOUNGER SWITCH DRIVERS. 235
FIGURE 8.45. EVENTS PER MILE FOR LEASED AND PRIVATE VEHICLE DRIVING IN THE FIRST 10 HOURS OF DRIVING FOR A MATCHED SET OF 6 SWITCH DRIVERS. 235
FIGURE 8.46. EVENTS PER MILE FOR LEASED AND PRIVATE VEHICLE DRIVING IN THE FIRST 10 HOURS OF DRIVING FOR A MATCHED SET OF 4 YOUNGER SWITCH DRIVERS. 236
FIGURE 9.1. FREQUENCY OF LEAD- AND FOLLOWING-VEHICLE EVENTS BY LEVEL OF SEVERITY. 241
FIGURE 9.2. THE RATE OF LEAD- AND FOLLOWING-VEHICLE CONFLICTS BY LEVEL OF SEVERITY. 242
FIGURE 10.1. RATE PER MVMT FOR SV STRIKING BY AGE AND LANE CHANGE MANEUVER FOR INCIDENTS. 271
FIGURE 10.2. RATE PER MVMT FOR SV STRIKING BY AGE AND LANE CHANGE MANEUVER FOR NEAR-CRASHES. 271
FIGURE 10.3. RATE PER MVMT FOR SV STRIKING BY AGE AND BY GENDER BY LANE CHANGE MANEUVER FOR INCIDENTS. 272
FIGURE 10.4. RATE PER MVMT FOR SV STRIKING BY AGE AND GENDER BY LANE CHANGE MANEUVER FOR NEAR-CRASHES. 273
FIGURE 10.5. RATE PER MVMT FOR STRUCK SV BY AGE AND LANE CHANGE MANEUVER FOR INCIDENTS. 274
FIGURE 10.6. RATE PER MVMT FOR STRUCK SV BY AGE AND LANE CHANGE MANEUVER FOR NEAR-CRASHES. 274
FIGURE 10.7. RATE PER MVMT FOR STRUCK SV BY AGE AND GENDER BY LANE CHANGE MANEUVER FOR INCIDENTS. 275
FIGURE 10.8. RATE PER MVMT FOR STRUCK SV BY AGE AND GENDER BY LANE CHANGE MANEUVER FOR NEAR-CRASHES. 275
FIGURE 10.9. RANGE AND RANGE-RATE VALUES FOR CUTOFF LANE CHANGE CONFLICTS. 280
FIGURE 11.1. FREQUENCY OF CRASHES, NEAR-CRASHES, AND INCIDENTS WHERE INATTENTION WAS A FACTOR. 286
FIGURE 11.2. PERCENTAGE OF LEAD-VEHICLE CRASHES, NEAR-CRASHES, AND INCIDENTS FOR WHICH INATTENTION WAS LISTED AS A CONTRIBUTING FACTOR. 287
FIGURE 11.3. THE FREQUENCY OF INATTENTION CATEGORIES FOR CRASHES AND NEAR-CRASHES. 288
FIGURE 11.4. THE PERCENTAGE OF INATTENTION CATEGORIES FOR CRASHES AND NEAR-CRASHES. 288
FIGURE 11.5. FREQUENCY OF INATTENTION-RELATED LEAD-VEHICLE CONFLICT INCIDENTS BY TYPE OF INATTENTION OR COMBINATION OF TYPES OF INATTENTION. 289
FIGURE 11.6. FREQUENCY OF CRASHES BY DRIVER ATTENTION AND LEAD-VEHICLE KINEMATIC SCENARIO. 290
FIGURE 11.7. FREQUENCY OF NEAR-CRASHES BY DRIVER ATTENTION LEVEL AND LV SCENARIO. 291
FIGURE 11.8. FREQUENCY OF INCIDENTS BY DRIVER ATTENTION LEVEL AND LV SCENARIO. 291
FIGURE 11.9. TOTAL FREQUENCY OF SECONDARY TASK INATTENTION SOURCES FOR LEAD-VEHICLE EVENTS. 293
FIGURE 11.10. TOTAL FREQUENCY OF SECONDARY TASK TYPE BY SEVERITY. 294
FIGURE 11.11. CONTRIBUTION OF DROWSINESS TO THE OVERALL NUMBER OF INATTENTION-RELATED LEAD-VEHICLE CRASHES, NEAR-CRASHES, AND INCIDENTS. 295
FIGURE 11.12. FREQUENCY OF GLANCE LOCATIONS FOR LEAD-VEHICLE EVENTS CATEGORIZED AS DRIVING-RELATED INATTENTION. 296
FIGURE 11.13. FREQUENCY OF CRASH AND NEAR-CRASH EVENTS FOR WHICH A GLANCE TO A NONSPECIFIC LOCATION WAS PRESENT. 297

FIGURE 11.14. FREQUENCY OF *LV STOPPED >2 S* INATTENTION-RELATED EVENTS BY INATTENTION SOURCE AND EVENT SEVERITY. 298
FIGURE 11.15. FREQUENCY OF *LV STOPPED >2 S* INATTENTION-RELATED INCIDENTS BY INATTENTION. 298
FIGURE 11.16. FREQUENCY OF SECONDARY TASK INATTENTION SOURCES FOR *LV STOPPED >2 S* EVENTS BY EVENT SEVERITY. 299
FIGURE 11.17. FREQUENCY OF *LV STOPPED < 2 S* INATTENTION-RELATED CRASHES AND NEAR-CRASHES BY EVENT SEVERITY. 300
FIGURE 11.18. FREQUENCY OF *LV STOPPED < 2 S* INATTENTION-RELATED INCIDENTS BY INATTENTION SOURCE. 300
FIGURE 11.19. FREQUENCY OF SECONDARY TASK INATTENTION SOURCES FOR *LV STOPPED < 2 S* EVENTS BY EVENT SEVERITY. 301
FIGURE 11.20. FREQUENCY OF *LV DECELERATING* INATTENTION-RELATED NEAR-CRASHES. NOTE: THERE WERE NO CRASHES IN THIS CATEGORY. 302
FIGURE 11.21. FREQUENCY OF *LV DECELERATING INATTENTION*-RELATED INCIDENTS. 302
FIGURE 11.22. FREQUENCY OF SECONDARY TASK INATTENTION SOURCES FOR *LV DECELERATING* EVENTS BY EVENT SEVERITY. 303
FIGURE 11.23. FREQUENCY OF *LV MOVING AT SLOWER CONSTANT SPEED* INATTENTION-RELATED NEAR-CRASHES. NOTE: THERE WERE NO CRASHES IN THIS CATEGORY. 304
FIGURE 11.24. FREQUENCY OF *LV TRAVELING AT A SLOWER, CONSTANT SPEED INATTENTION*-RELATED INCIDENTS. 304
FIGURE 11.25. FREQUENCY OF SECONDARY TASK INATTENTION SOURCES FOR *LV MOVING AT SLOWER CONSTANT SPEED* EVENTS BY EVENT SEVERITY. 305
FIGURE 11.26. FREQUENCY OF LV EVENTS FOR WIRELESS DEVICES BY SCENARIO AND SEVERITY. 306
FIGURE 11.27. FREQUENCY OF LV EVENTS FOR PASSENGER-RELATED INATTENTION BY SCENARIO AND SEVERITY. 307
FIGURE 12.1. HEINRICH'S TRIANGLE SHOWING THE RELATIVE OCCURRENCE OF LEAD-VEHICLE CONFLICT EVENTS BY SEVERITY. 310
FIGURE 12.2. HEINRICH'S TRIANGLE SHOWING THE RELATIVE OCCURRENCE OF *LV STOPPED >2 S* EVENTS BY SEVERITY. 311
FIGURE 12.3. HEINRICH'S TRIANGLE SHOWING THE RELATIVE OCCURRENCE OF *LV STOPPED ≤ 2 S* EVENTS BY SEVERITY. 311
FIGURE 12.4. HEINRICH'S TRIANGLE SHOWING THE RELATIVE OCCURRENCE OF *LV DECELERATING* EVENTS BY SEVERITY. 312
FIGURE 12.5. HEINRICH'S TRIANGLE SHOWING THE RELATIVE OCCURRENCE OF *LV MOVING AT A SLOWER CONSTANT SPEED* EVENT BY SEVERITY. 313
FIGURE 12.6. HEINRICH'S TRIANGLE SHOWING THE RELATIVE OCCURRENCE OF *LV ACCELERATING* EVENTS BY SEVERITY. 313

LIST OF TABLES

Table RO.1. Driver age and gender distributions. xxxiii
Table RO.2. Actual miles driven during the study. xxxiii
Table RO.3. Number of crashes, near-crashes, and incidents for each conflict type. xxxvi
Table RO.4. Crash type by crash severity category. xxxvii
Table RO.5. Number and percentage of drivers involved in multiple events. xxxviii
Table RO.6. Frequencies for the 5 RE lead-vehicle scenarios by event severity. xlviii
Table RO.7. Frequencies for the 5 RE following-vehicle scenarios by event severity. xlviii
Table RO.8. Rates and Confidence Limits for each Lead-vehicle conflict severity category. lvi
Table RO.9. Catastrophic or major failure rates by sensor or subsystem. lvii
Table RO.10. Minor failure rates by sensor or subsystem. lviii
Table RO.11. "Confusion matrix" showing hits, misses, false alarms, and correct rejections for the classification of near-crash events based upon multivariate trigger criteria. lviii
Table 1.1. List of research goals organized by research category and subcategory. 3
Table 1.2. Severity levels for the 100-Car Study. 6
Table 2.1. Description of Sensor Components. 17
Table 2.2. Description of all tests and questionnaires administered to study participants. 18
Table 2.3. Dependent variables used as event triggers. 20
Table 2.4. Total number of events reduced for each severity level. 22
Table 2.5. Areas of data reduction, definition of the area, and examples. 24
Table 2.6. Percentage agreement with expert reductionists. 25
Table 2.7. Top-level 100-car study statistics. 26
Table 2.8. Summary of crashes. 27
Table 2.9. Ratios of non-police-reported to police reported crashes. 27
Table 2.10. Number and percentage of drivers involved in at least one of the various event types. 28
Table 2.11. Number and percentage of drivers involved in multiple events. 29
Table 2.12. Description of four outlier participants in terms of crash, near-crash, and incidents rates per miles traveled. 30
Table 3.1. Participant age and gender distributions. 31
Table 3.2. Comparison of age and gender distribution to project goals. 32
Table 3.3. Ethnic background of northern Virginia residents and 100-Car Study participants. 35
Table 3.4. Occupation categories of participants by gender. 36
Table 3.5. Actual miles driven during the study. 43
Table 5.1. Precipitating Factors Used to Classify 100-Car Study Events. 59
Table 5.2. Number of crashes, near-crashes, and incidents for each conflict type. 62
Table 5.3. Crash type by crash severity category. 63
Table 5.4. Stimulus response time and crash times for following-vehicle level I, II, III crashes. 93
Table 5.5. Average stimulus response time and crash times for following-vehicle level I, II, III crashes. 94
Table 5.6. Omitted epics in SRT and crash time calculations for following-vehicle level I, II, III crashes. 94
Table 5.7. Factor Type Categories. 124
Table 5.8. Comparison of crash types for 100-Car Crashes, Near-crashes, Police-Reported Crashes, and GES Crashes. 125
Table 5.9. Pre-event maneuver, crash, and near-crash. 126
Table 5.10. Drivers' avoidance maneuver, crash, and near-crash. 127
Table 5.11. Driver's willful behavior, crash, and near-crash. 127
Table 5.12. Driver impairments, crash, and near-crash. 128
Table 5.13. Driver proficiency error, crash, and near-crash. 128
Table 5.14. Hands on wheel, crash, and near-crash. 128
Table 5.15. Inattention to forward roadway, crash, and near-crash. 130
Table 5.16. Visual obstruction, crash, and near-crash. 131
Table 5.17. Surface condition, crash, and near-crash. 131
Table 5.18. Relation to junction, crash, and near-crash. 132
Table 5.19. Road alignment/road profile, crash, and near-crash. 133
Table 5.20. Locality, crash, and near-crash. 133
Table 5.21. Lighting, crash, and near-crash. 134
Table 5.22. Weather, crash, and near-crash. 134

Table 5.23. Trafficway flow, crash, and near-crash....135
Table 5.24. Traffic density, crash, and near-crash....135
Table 5.25. Traffic control device, crash and near-crash....136
Table 5.26. Vehicle factors, crash and near-crash....136
Table 5.27. Infrastructure, crash and near-crash....137
Table 6.1. Percentage of hits and false alarms for each boundary model for a conflict with a lead vehicle....142
Table 6.2. Number of lead-vehicle crashes detected by each crash boundary based upon the number of seconds prior to the crash....147
Table 6.3. Percentage of hits and false alarms for each boundary model for a conflict with a following vehicle....147
Table 6.4. The number and percentage of events that would be correctly classified with the addition of a speed threshold for each boundary equation for lead-vehicle conflicts....153
Table 6.5. The number and percentage of events that would be correctly classified with the addition of a deceleration threshold for each boundary equation for lead-vehicle conflicts....155
Table 8.1. Example data matrix for calculating risk ratio. A, B, C, and D are the number of cases satisfying each criteria....187
Table 8.2. Mean number of events per vehicle per time period for leased and private vehicles; calculated relative risk for those time periods; 95th percentage confidence intervals based on RR....193
Table 8.3. Mean number of events per vehicle per time period for leased and private vehicles; calculated relative risk for those time period. The baseline leased vehicle mean is now set equal to the private vehicle baseline....194
Table 8.4. Mean number of events per vehicle per time period for leased and private vehicles; calculated relative risk for those time periods; 95th percentage confidence intervals based on RR....196
Table 8.5. Number of leased and private vehicles available for weeks 1-50 for younger drivers only....199
Table 8.6. Number of younger drivers of each age for leased and private vehicles....200
Table 8.7. Mean number of events for younger drivers only per time period for leased and private vehicles; calculated RR for those time periods. Leased vehicle values do not include leased portion of switch drivers....201
Table 8.8. Mean number of near-crashes for younger drivers only per time period for leased and private vehicles; calculated RR for those time periods. Leased vehicle values do not include leased portion of switch drivers....204
Table 8.9. Mean number of crashes for younger drivers only per time period for leased and private vehicles; calculated RR for those time periods. Leased vehicle values do not include leased portion of switch drivers....207
Table 8.10. Mean number of events for per time period for leased and private vehicles; calculated RR for those time periods....211
Table 8.11. Mean number of events per time period for younger driver leased and private vehicles; calculated RR for those time periods....213
Table 8.12. Mean number of near-crashes per time period for younger driver leased and private vehicles; calculated RR for those time periods....215
Table 8.13. Statistics for a matched set of all switch drivers for weeks 1-4....217
Table 8.14. Statistics for a matched set of all younger switch drivers for weeks 1-4....219
Table 8.15. Statistics for a matched set of all older switch drivers for weeks 1-4....221
Table 8.16. Mean number of events per vehicle per mile for leased and private vehicles; calculated relative risk for those time period....225
Table 8.17. Mean number of events per mile for younger drivers only for leased and private vehicles; calculated RR for those time periods. Leased vehicle values do not include leased portion of switch drivers....227
Table 8.18. Mean number of events for per mile for leased and private vehicles; calculated RR for those time periods....230
Table 8.19. Events per mile statistics for a matched set of all switch drivers for weeks 1-4....232
Table 9.1. Frequencies for the 5 RE lead-vehicle scenarios by event severity....243
Table 9.2. Rate of events per MVMT for the 5 RE lead-vehicle scenarios by event severity....243
Table 9.3. Rate of events per MVMT for the 5 RE lead-vehicle scenarios for each age group by event severity....244
Table 9.4. Rate of events per MVMT for the 5 RE lead-vehicle scenarios by severity, age group, and gender....245
Table 9.5. Frequencies for the 5 RE following-vehicle scenarios by event severity....246
Table 9.6. Rate of events per MVMT for the 5 RE following-vehicle scenarios by event severity....246
Table 9.7. Rate of events per MVMT for the 5 RE following-vehicle scenarios for each age group by event severity....247

Table 9.8. Rate of events per MVMT for the 5 RE following-vehicle scenarios by severity, age group, and gender 248
Table 9.9. Rate of events per MVMT for the 5 RE lead-vehicle scenarios for each vehicle speed category by event severity.....250
Table 9.10. Rate of events per MVMT for the 5 RE lead-vehicle scenarios for each vehicle headway category by event severity. 251
Table 9.11. Rate of events per MVMT for the 5 RE following-vehicle scenarios for each vehicle speed category by event severity 252
Table 9.12. Rate of events per MVMT for the 5 RE following-vehicle scenarios for each vehicle headway category by event severity 253
Table 9.13. Rate of events per MVMT for the 5 RE lead-vehicle scenarios for each driver contributing factor by event severity 255
Table 9.14. Rate of events per MVMT for the 5 RE lead-vehicle scenarios for each environmental and roadway contributing factor by event severity.....256
Table 9.15. Rate of events per MVMT for the 5 RE following-vehicle scenarios for each driver contributing factor by event severity 258
Table 9.16. Rate of events per MVMT for the 5 RE following-vehicle scenarios for each environmental and roadway contributing factor by event severity.....259
Table 9.17. Rate of events per MVMT for the 5 RE lead-vehicle scenarios for each corrective action category. ...261
Table 9.18. Rate of events per MVMT for RE lead-vehicle scenarios by age and gender for each corrective action category.....261
Table 9.19. Rate of events per MVMT for the 5 RE following-vehicle scenarios for each corrective action category. 262
Table 9.20. Rate of events per MVMT for RE following-vehicle scenarios by age and gender for each corrective action category. 263
Table 10.1. Frequencies of conflicts with lead-vehicles and conflict with following-vehicle conflicts associated with different lane change maneuvers.....269
Table 10.2. Rates per MVMT of conflicts with lead vehicles and conflicts with following vehicles associated with different lane change maneuvers.....270
Table 10.3. Rate of incidents per MVMT for SV struck and SV striking environmental and roadway factors. 277
Table 10.4. Rate of near-crashes per MVMT for SV struck and SV striking environmental and roadway factors...277
Table 10.5. Rate of incidents per MVMT for SV struck and SV striking driver factors. 278
Table 10.6. Rate of near-crashes per MVMT for SV struck and SV striking driver factors.....278
Table 10.7. Lane change conflict frequencies by direction and type for side sensor-equipped vehicles.....279
Table 12.1. Lead-vehicle conflicts modeled using Poisson distribution to determine event rate per MVMT, variance, and 95 percent confidence intervals. 310
Table 12.2. *LV stopped >2 seconds* data modeled using Poisson distribution to determine event rate per MVMT, variance, and 95 percent confidence intervals.....311
Table 12.3. *LV stopped ≤ 2 seconds* data modeled using Poisson distribution to determine event rate per MVMT, variance, and 95 percent confidence intervals.....312
Table 12.4. *LV decelerating* data modeled using Poisson distribution to determine event rate per MVMT, variance, and 95 percent confidence intervals. 312
Table 12.5. *LV moving at slower constant speed* data modeled using Poisson distribution to determine event rate per MVMT, variance, and 95 percent confidence intervals. 313
Table 12.6. *LV accelerating* data modeled using Poisson distribution to determine event rate per MVMT, variance, and 95 percent confidence intervals 314
Table 13.1. Number of vehicles on the road for each month of the study. 318
Table 13.2. Causal factors for catastrophic and major failures.....321
Table 13.3. Causal factors for minor failures. 321
Table 13.4. Percentage of failures for each month of the study.....322
Table 13.5. Catastrophic or major failure rates by sensor or subsystem.....323
Table 13.6. Minor failure rates by sensor or subsystem. 323
Table 13.7. Sensor or subsystem benefits in terms of valid events captured.....326
Table 13.8. Sensor or subsystem benefits in terms of proportion of invalid events. 328
Table 13.9. Trigger characteristics for the 100-Car Study.....331
Table 13.10. Confusion matrix for passenger cars.....337
Table 13.11. Trigger Confusion matrix for SUVersus.....337

Table 13.12. Confusion matrix for different event types. 338
Table 13.13. Confusion matrix for different event severities. 339
Table 13.14. Confusion matrix for conflict with lead-vehicle events. 339
Table 13.15. Confusion matrix for conflict with following vehicle events. 339
Table 13.16. Confusion matrix for conflict with lead-vehicle events, equal event probability. 340
Table 13.17. Confusion matrix for conflict with lead-vehicle events, 0.7 invalid event probability. 340
Table 13.18. Confusion matrix for conflict with lead-vehicle events, 0.9 invalid event probability. 340
Table 13.19. Valid events remaining as a function of deceleration threshold selected for pre-discriminant analysis filter. 342
Table 13.20. Discriminant analysis classification results as a function of expected probability and percentile threshold. 342
Table 13.21. Confusion matrix for conflict with lead-vehicle events, 40th percentile longitudinal acceleration, probabilities based on sample. 343
Table 13.22. Confusion matrix for conflict with lead-vehicle events using the logistic regression approach. 344

GLOSSARY OF TERMS

ANOVA – Analysis of variance.

Additional driver – Family or friends of the primary driver who drove the subject's vehicle and were not involved with the in-processing.

Associative Factors – Any environmental or vehicular factor where direct causation to crashes, near-crashes, or incidents is not possible to attain but correlation may be determined.

Backing crash – A crash that occurs while the driver's vehicle is in reverse gear.

Chase vehicle – Vehicle designated for locating (through GPS or other means) and downloading data from subject vehicles.

Contributing factors – Any circumstance that leads up to or has an impact on the outcome of the event. This term encompasses driver proficiency, willful behavior, roadway infrastructure, distraction, vehicle contributing factors, and visual obstructions.

Crash – Any contact with an object, either moving or fixed, at any speed in which kinetic energy is measurably transferred or dissipated. Includes other vehicles, roadside barriers, objects on or off of the roadway, pedestrians, cyclists, or animals.

Crash-Relevant Event – Any circumstance that requires a crash avoidance response on the part of the subject vehicle, any other vehicle, pedestrian, cyclist, or animal that is less severe than a rapid evasive maneuver (as defined above), but greater in severity than a "normal maneuver" to avoid a crash. A crash avoidance response can include braking, steering, accelerating, or any combination of control inputs. A "normal maneuver" for the subject vehicle is defined as a control input that falls outside of the 95 percent confidence limit for control input as measured for the same subject.

Conflict Type – All crashes, near-crashes, crash-relevant conflicts and proximity conflicts were categorized based on the initial conflict that lead to the crash that occurred or would have occurred in the case of near-crashes and incidents. There were 20 types of conflicts used which are as follows: conflict with lead vehicle, following vehicle, oncoming traffic, vehicle in adjacent lane, merging vehicle, vehicle turning across subject vehicle path (same direction), vehicle turning across subject vehicle path (opposite direction), vehicle turning into subject vehicle path (same direction), vehicle turning into subject vehicle path (opposite direction), vehicle moving across subject vehicle path (through intersection), parked vehicle, pedestrian, pedalcyclist, animal, obstacle/object in roadway, single vehicle conflict, other, no known conflict, unknown conflict. This list was primarily NASS GES Accident Types.

DAS – Data Acquisition System.

Driver Impairment – The driver's behavior, judgment, or driving ability is altered or hindered. Includes drowsiness, use of drugs or alcohol, illness, lack of, or incorrect use of medication, or disability.

Driver Proficiency – Whether the individual's driving skills, abilities, or knowledge are inadequate. This specifically refers to whether the driver appeared to be aware of specific traffic laws (i.e., no U-turn), whether the driver was incompetent to safely perform a driving maneuver (i.e., check for traffic before pulling out on a roadway), unaware of the vehicle's turning radius, or performs driving maneuvers under the incorrect assumption that it is safe, (i.e., drives over a concrete median).

Driver-Related Inattention to the Forward Roadway – Inattention due to a necessary and acceptable driving task where the subject is required to shift attention away from the forward roadway. (e.g., checking blind spots, center mirror, instrument panel).

Driver Reaction – The evasive maneuver performed in response to the precipitating event.

Driver Seat Belt Use – Variable indicating if the subject is wearing a seat belt during an event.

EDR – Electronic data recorder.

Epoch – Typically, a 90-second period of time around one or more triggers in the data; can include one or more events.

Event – a term referring to all crashes, near-crashes, and incidents. The "event" begins at the onset of the precipitating factor and ends after the evasive maneuver.

Event Nature – Classification of the type of conflict occurring in the event (e.g., Conflict with lead vehicle, Conflict with vehicle in adjacent lane).

Event Severity – Classification of the level of harm or damage resulting from an event. The 5 levels were crash, near-crash, crash-relevant, proximity, and nonconflict.

FARS – Fatality Analysis Reporting System.

FOV – Field of view.

FV – Following vehicle**.**

GPS – Global Positioning System **–** used by reductionists to locate participant vehicle for information on an event.

Inattention Event – Any event where drowsiness, driver-related inattention to the forward roadway, driver secondary tasks, or nonspecific eyeglance away from the forward roadway were identified as a contributing factors to the event.

Incident – Encompasses the event severities of crash-relevant conflicts and proximity conflicts.

IVI – Intelligent Vehicle Initiative**.**

IR LEDs – Infrared light emitting diode.

Invalid Trigger – Any instance where a pre-specified signature in the driving performance data stream is observed but no safety-relevant event is present. See Appendix B for a more complete definition of triggers.

LV – Lead vehicle.

MVMT – Million vehicle miles traveled.

NHTSA – National Highway Traffic Safety Administration.

Naturalistic – Unobtrusive observation; observation of behavior taking place in its natural setting.

Near-crash – Any circumstance that requires a rapid, evasive maneuver by the subject vehicle, or any other vehicle, pedestrian, cyclist, or animal to avoid a crash. A rapid, evasive maneuver is defined as a steering, braking, accelerating, or any combination of control inputs that approaches the limits of the vehicle capabilities.

Non-Conflict – Any incident that increases the level of risk associated with driving, but does not result in a crash, near-crash, or incident as defined above. Examples include driver control error without proximal hazards being present, driver judgment error such as unsafe tailgating or excessive speed, or cases in which drivers are visually distracted to an unsafe level.

Non-Subject Conflict – Any incident that gets captured on video, crash-relevant, near-crash, or crash, that does not involve the subject driver. Labeled as a non-subject conflict but data reduction was not completed.

ORD – Observer Rating of Drowsiness; measured on a scale from 0 to 100 in increasing severity of drowsiness. Based on Wierwille and Ellsworth, 1994.

Precipitating factor – The driver behavior or state of the environment that initiates the crash, near-crash, or incident and the subsequent sequence of actions that result in an incident, near-crash, or crash.

Proximity event – Any circumstance resulting in extraordinarily close proximity of the subject vehicle to any other vehicle, pedestrian, cyclist, animal, or fixed object where, due to apparent unawareness on the part of the driver(s), pedestrians, cyclists or animals, there is no avoidance maneuver or response. Extraordinarily close proximity is defined as a clear case where the absence of an avoidance maneuver or response is inappropriate for the driving circumstances (including speed, sight distance, etc.).

Pre-Incident Maneuver – The maneuver that our driver was performing immediately prior to the event. The importance of this is to record what our driver was doing before the precipitating event occurred.

Secondary Task – Task, unrelated to driving, which requires subjects to divert attentional resources from the driving task, e.g., talking on the cell phone, talking to passenger, eating, etc.

Rear-end striking – Refers to the subject vehicle striking a lead vehicle.

Rear-end struck - Refers to the subject vehicle being struck by a following vehicle.

Sideswipe – Refers to either a vehicle in the adjacent lane changing lanes into the subject vehicle or the subject vehicle changing lanes into a vehicle in the adjacent lane.

SV – Subject vehicle.

Trigger/Trigger Criteria – A signature in the data stream that, when exceeded, 90 seconds of video data (60 seconds prior and 30 seconds after the data excedence) and the corresponding driving performance data are copied and saved to a database. Trained data reductionists assess these segments of video and driving performance data to determine whether this segment of data contains a safety-relevant conflict (i.e., crash, near-crash, or incident) or not. Examples of triggers include a driver braking at 0.76 g longitudinal deceleration or swerving around an obstacle obtaining a 0.8 g lateral acceleration. For a more complete description of triggers, see Appendix B.

USDOT – United States Department of Transportation.

Valid Event or Valid Trigger – Those events where a specific signature in the data stream was identified, viewed by a data reductionist, and deemed to contain a safety-relevant scenario. Data reductionists record all relevant variables and store this data in the 100-Car Database.

Vehicle Run-Off-Road – Describes a situation when the subject vehicle departs the roadway.

VDOT – Virginia Department of Transportation.

Virginia Tech Motor Pool – An extension of the Virginia Tech Office of Transportation.

VTTI – Virginia Tech Transportation Institute.

Visual Obstruction – This variable refers to glare, weather, or an object obstructing the view of the driver that impacts the event in any way.

Willful Behavior – The driver knowingly and purposefully drives in an unsafe or inappropriate manner. Includes aggressive driving, purposeful violation of traffic laws, use of vehicle for improper purposes (i.e., intimidation).

EXECUTIVE SUMMARY

The 100-Car Naturalistic Driving Study is the first instrumented vehicle study undertaken with the primary purpose of collecting large-scale naturalistic driving data. Drivers were given no special instructions, no experimenter was present, and the data collection instrumentation was unobtrusive. In addition, the majority of the drivers drove their own vehicles (78 out of 100 vehicles). As described throughout this document, there is every indication that the drivers rapidly disregarded the presence of the instrumentation. Thus, the resulting database contains many extreme cases of driving behavior and performance, including severe drowsiness, impairment, judgment error, risk taking, willingness to engage in secondary tasks, aggressive driving, and traffic violation (just to name a few) that have been heretofore greatly attenuated by other empirical techniques.

Since the study was the first of its kind, new techniques had to be created and existing methods modified to make the study successful. The data collection effort resulted in the following dataset contents:

- Approximately 2,000,000 vehicle miles of driving.
- Almost 43,000 hours of data.
- 241 primary and secondary driver participants.
- 12 to 13 month data collection period for each vehicle; 18 month total data collection period.
- Five channels of video and many vehicle state and kinematic variables.

An "event" database was created, similar in classification structure to an epidemiological crash database, but with video and electronic driver and vehicle performance data appended to it. The events in this case are crashes, near-crashes, and other "incidents" that represent less severe conflicts. This approach allows the video and electronic data to be replayed multiple times and at varying frame rates in order to fully understand the nature of the event. This approach allows the classification of the following:

- Pre-event maneuver.
- Precipitating factor.
- Event type.
- Contributing factors.
- Associative factors.
- Avoidance maneuver.

The scope of the current project specified 10 initial, high priority objectives or goals addressed through the initial analysis of the event database. This report addresses the first 9 of these 10 goals, which include:

Goal 1: Characterization of crashes, near-crashes, and incidents for the 100-Car study
Goal 2: Quantification of near-crash events
Goal 3: Characterization of driver inattention
Goal 4: Driver behavior over time
Goal 5: Rear-end conflict causal factors and dynamic conditions
Goal 6: Lane change causal factors and dynamic conditions
Goal 7: Inattention for rear-end lead-vehicle scenarios
Goal 8: Characterize the rear-end scenarios in relation to Heinrich's Triangle
Goal 9: Evaluate performance of hardware, sensors, and the data collection system.
Goal 10: Evaluate the data reduction plan, triggering methods, and data analysis

Some of the most important findings addressed as part of the high priority goals analyzed for this report are presented below:

- This study allowed, perhaps for the first time, the capture of crash and collision events that included minor, non-property-damage contact. These low severity collisions provide very valuable information and occur much more frequently than more severe crashes. As a result, crash/collision-involvement was much higher than expected in that 82 total crashes/collisions were reported in this study, while only 15 of these crashes were reported to the police. For urban/suburban settings, this suggests that total crash/collision involvement may be over five times higher than police-reported crashes.

- Almost 80 percent of all crashes and 65 percent of all near-crashes involved the driver looking away from the forward roadway just prior to the onset of the conflict. Prior estimates related to "distraction" as a contributing factor have been in the range of 25 percent.

- Inattention, which was operationally defined as including: (1) secondary task distraction; (2) driving-related inattention to the forward roadway (e.g., blind spot checks); (3) moderate to extreme drowsiness; and (4) other non-driving-related eyeglances, was a contributing factor for 93 percent of the conflict with lead-vehicle crashes and minor collisions. In 86 percent of the lead-vehicle crashes/collisions, the headway at the onset of the event was greater than 2.0 seconds.

- For scenarios involving conflict with a lead vehicle, the most frequent cases of lower severity conflicts (i.e., incidents and near-crashes) occurred in lead-vehicle moving scenarios, while 100 percent of the crashes (14 total) occurred when the lead vehicle was stopped. This indicates that drivers have sufficient awareness and ability to perform evasive maneuvers when closing rates are lower and/or expectancies about the flow of traffic are not violated.

- The rate of inattention-related crash and near-crash events decreases dramatically with age, with the rate being as much as four times higher for the 18-to-20 age group relative to some of the older driver groups (i.e., 35 and up).

- The use of hand-held wireless devices (primarily cell phones but including a small amount of PDA use) was associated with the highest frequency of secondary task distraction-related events. This was true for both events of lower severity (i.e., incidents) and for events of higher severity (i.e., near-crashes). Wireless devices were also among the categories associated with the highest frequencies of crashes and minor collisions, along with looking at/reaching for an object in vehicle and passenger-related secondary tasks.

- Drowsiness also appears to affect crashes and collisions at much higher rates than is reported using existing crash databases. Drowsiness was a contributing factor in 12 percent of all crashes and 10 percent of near-crashes, while most current database estimates place drowsiness-related crashes at approximately 2 to 4 percent of total crashes.

- The lead-vehicle crash and near-crash data clearly shows that development of purely quantitative near-crash criteria (i.e., not requiring at least some degree of verification by a human analyst) is not currently feasible. A primary reason for this was that vehicle kinematics associated with near-crashes were virtually identical to common driving situations that were not indicative of crash risk. Thus, qualitative and quantitative criteria are dependent upon one another to some degree. Fortunately, advances in digital video compression and storage technology, and the advancement of data reduction software, have made video verification feasible for large numbers of events.

- Results from the analysis investigating driver adaptation to instrumented vehicles indicate that even when the same driver was switched from a private vehicle to a leased vehicle, there were still more events per mile in the leased vehicle than in the private vehicle. If there was an effect of adaptation, it was extinguished before the first week of driving was completed. In addition, drivers appeared to adapt to the presence of the unobtrusive instrumentation within the first hour of driving.

In addition to the 10 high-priority goals addressed as part of this report, there are three additional research contracts in place to perform further data reduction and analysis efforts for the purpose of addressing another 8 goals. There is also considerable interest in using the data for even more purposes from researchers in several disciplines. Progressing toward this potential for a multipurpose, highly flexible and adaptable tool for driving safety may be the most important aspect of this study.

The naturalistic approach fills a void in our existing driving safety research methods. Specifically, it provides much greater information regarding the pre-crash and crash events than is currently available, even after a detailed crash investigation. Furthermore, the data provides much greater external validity relative to the larger context of driving than do empirical methods such as test tracks or simulators.

Despite the massive scope of the current effort, it was designed to also serve as a pilot to a much larger future study. From an epidemiological viewpoint, the study was small with the presence of 15 police-reported and 82 total crashes and minor collisions. Furthermore, drivers were represented from only one area of the country (Northern Virginia/Washington, DC, metro area).

One purpose of a larger-scale study would be to have a statistically representative sample of crashes (perhaps 2,000) and a more representative subject/environment sample.

Since a primary purpose of the 100-Car Study was to serve as a pilot for a larger-scale study (e.g., 5,000-car study), a goal was to evaluate the process and results of the 100-Car Study to assess the feasibility of such an undertaking. Based upon the results of the evaluations conducted, it is believed that a large-scale database would be an enormous asset and would be used by transportation researchers to address many transportation safety problems. Such an undertaking would allow researchers to gain insight and understanding into a wide array of driving behavior issues and potentially serve as a basis for decision making and program development in both the public and private sectors. This belief is based upon the robustness of these pilot results and the anticipation that these data will continue to be analyzed and the results made available from a variety of researchers and research organizations. Clearly, a large-scale, nationally-representative study, that includes a statistically significant number of police–reported crashes, would provide tremendous insight into issues that have eluded the highway safety community for many years.

REPORT OVERVIEW

INTRODUCTION

The 100-Car Naturalistic Driving Study is the first instrumented vehicle study designed to collect a large volume of naturalistic driving data for a large number of drivers over an extended period of time. The Virginia Tech Transportation Institute (VTTI) installed instruments and sensors in 100 vehicles that were then driven as ordinary vehicles by ordinary drivers for one year. Drivers were given no special instructions, no experimenter was present, and the data collection system was unobtrusive. In addition, drivers' own vehicles were instrumented for 78 out of the 100 vehicles. Drivers apparently adapted rapidly to the instrumentation, probably within the first hour. The resulting database contains many extreme cases of driving behavior and performance, including severe drowsiness, impairment, judgment error, risk taking, willingness to engage in secondary tasks, aggressive driving, and traffic violation (just to name a few) that have been difficult to examine using other techniques.

As with any innovative new research method, new techniques had to be developed and existing methods modified. The resulting dataset contained:

- approximately 2,000,000 vehicle-miles of driving;
- almost 43,000 hours of data;
- data on 241 primary and secondary drivers;
- a 12- to 13-month data collection period for each vehicle; and
- five channels of video and numerous vehicle state and kinematic variables for any given point in time.

Despite the apparent large scope of the current effort, the study was also designed to also serve as a pilot to a much larger future study. From an epidemiological viewpoint, the study was small (15 police-reported and 82 total crashes, including minor collisions). Furthermore, drivers from only one area of the country were represented (the Northern Virginia/Washington, DC, metro area). One purpose of a larger-scale study would be to have a statistically representative sample of perhaps 2,000 crashes as well as a more representative subject/environment sample.

Figure RO.1 shows how the approach used in the 100-Car Naturalistic Driving Study can fill in the gaps from existing driving safety research methods. On one hand, the 100-Car Study approach provides much greater information regarding the pre-crash and crash events than is currently available in crash databases, even those containing detailed crash investigation variables. On the other hand, the data provides much more naturalistic driving data than data obtained on test tracks or in simulators.

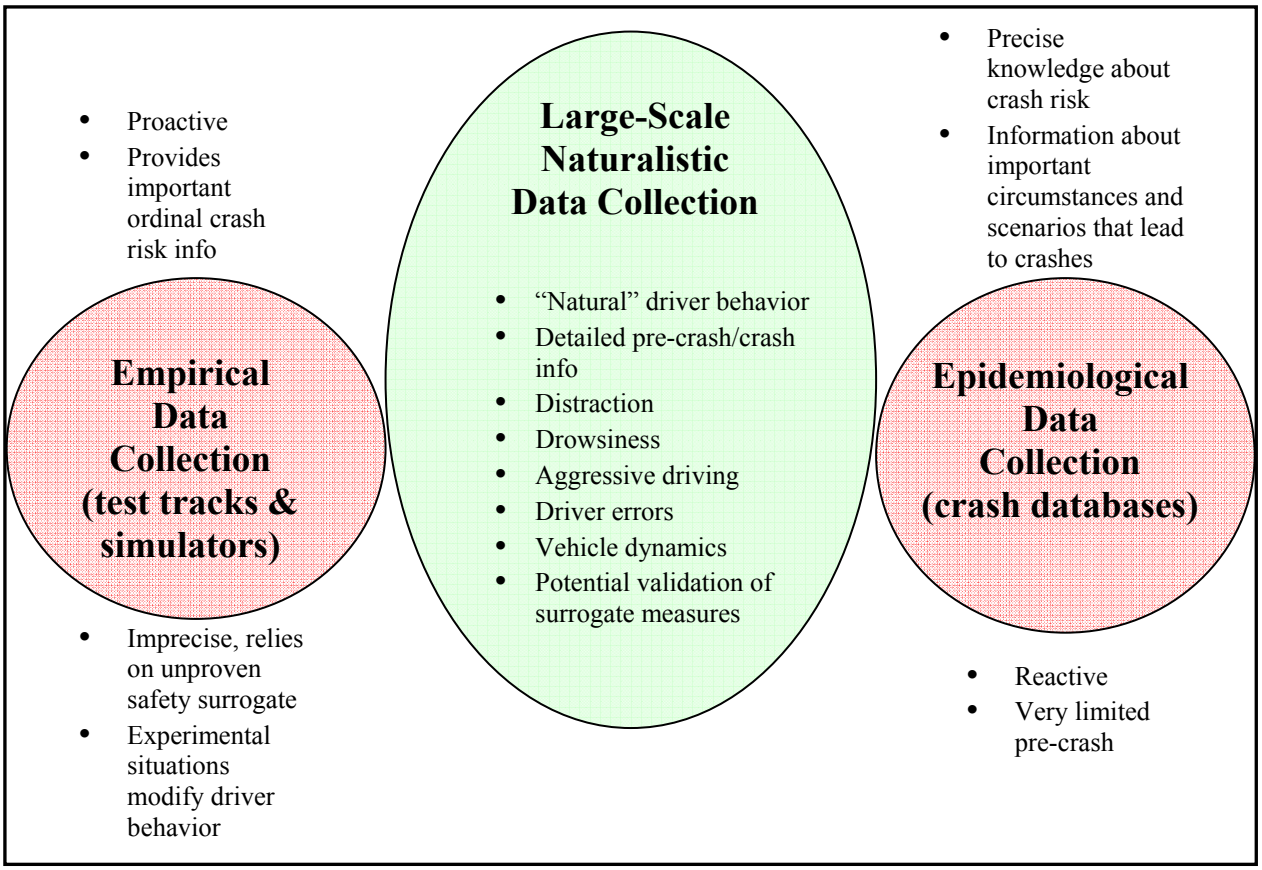


Figure RO.1. The relationship between empirical, naturalistic, and epidemiological methods in driving safety research.

Precise analysis of the events leading up to a crash or near-crash is possible with the 100-Car Study dataset since both video and electronic sensor data are available. In contrast, police reports and crash investigations rely on eyewitness accounts, and such data have been shown to be limited in accuracy. For example, drivers often do not remember specific, rapidly-occurring events as a crash or near-crash scenario unfolds. There are also cases in which the drivers or passengers are in shock or injured in a crash event, or in which they are trying to hide the details of what occurred (due either to embarrassment or fear of prosecution/litigation).

The 100-Car Study also marks the first time that detailed information on near-crash events has been collected. Near-crashes have two important advantages over crashes. First, they occur much more frequently (e.g., 15 times more often than crashes). Second, every near-crash event demonstrates a driver successfully performing an evasive maneuver. This may provide additional insight into effective defensive driving techniques and factors, as well as insight into potential countermeasures for these driving situations.

Unlike test track and simulator studies, naturalistic studies consider the larger context of driving. Furthermore, as demonstrated repeatedly in the 100-Car Study, the absence of an experimenter avoids potential modification of the driver's performance and behavior that may occur when a driver is directly observed.

One of the more notable contributions of the 100-Car Study is the creation of an "event" database. This database is similar in classification structure to an epidemiological crash database, but with video, driver, and vehicle data appended. The video and electronic data can be replayed multiple times and at varying frame rates in order to fully understand the nature of each event. The events in the dataset are crashes, near-crashes, and other "incidents" that represent less severe conflicts. Examples of reduced variables include:

- pre-event maneuver;
- precipitating factor;
- event type;
- contributing factors;
- associative factors; and
- avoidance maneuver.

One real advantage to this naturalistic approach is that the video allows direct viewing of all of the pre-event and during-event parameters, including the pre-event driver behaviors such as distraction, drowsiness, and error. In addition, this approach allows the precise calculation of parameters such as vehicle speed, vehicle headway, time-to-collision, and driver reaction time.

The resulting database should be useful for a variety of investigations for the next several years. In addition, the initial event database described above can be enhanced over time, since all of the video and electronic data for the entire study have been archived. The current project specified 10 objectives or *goals* to be addressed through the initial analysis of the event database. This report addresses the first 9 of these 10 goals. At the time this report is being written, three additional data reduction and analysis efforts are underway, and there is also considerable interest in using the data for additional research questions. The creation and improvement of this multi-purpose, flexible, and adaptable tool for driving safety may be one of the most important contributions of this study.

METHOD

Instrumentation

The 100-Car Study instrumentation package was engineered by VTTI to be rugged, durable, expandable, and unobtrusive. The system was the seventh generation of hardware and software that has been developed over the past 15 years. Previous iterations of the system have been deployed for a variety of traffic safety purposes. The system consisted of a Pentium-based computer that received and stored data from a network of sensors distributed around the vehicle. Data were stored on the system's hard drive, which could store several weeks of driving data before it needed to be downloaded.

Each of the sensing subsystems within a vehicle was independent, so that any failures were constrained to a single sensor type. Sensors included a box to obtain data from the vehicle network, an accelerometer box for longitudinal and lateral acceleration, a system to provide information on distance to lead and following vehicles, a system to detect conflicts with vehicles to either side of the subject vehicle, an incident box to allow drivers to flag incidents for the research team, a video-based lane tracking system to measure lane keeping behavior, and video to validate any sensor-based findings. The video subsystem was particularly important as it provided a continuous window into the happenings in and around the vehicle. There were 5

camera views monitoring the driver's face and driver's side view of the road, the forward road view, the rear road view, the passenger side road view, and an over-the-shoulder view for the driver's hands and surrounding areas. The video system was digital, with software-controllable video compression capability. This feature allowed synchronization, simultaneous display, and efficient archiving and retrieval of 100-Car Study data. A frame of compressed 100-Car Study video data is shown in Figure RO.2.

Figure RO.2. A compressed video image from the 100-Car Study data. The driver's face (upper left quadrant) has been distorted to protect the driver's identity. The lower right quadrant is split between the left-side view (top) and the rear view (bottom).

The 100-Car Study system had other capabilities that provided the research team with additional important information. These capabilities included automatic collision notification to inform the research team of possible collisions; cellular communications used by the research team to determine system status and vehicle position; system initialization equipment to automatically control system status; and a GPS positioning subsystem to collect information on vehicle position. The GPS positioning subsystem was 1 of 10 used in conjunction with the cellular communication subsystems to track and locate vehicles for repair and data downloading.

The main Data Acquisition System (DAS) unit was mounted under the package shelf for the sedans and behind the rear seat in the SUVs (Figures RO.3 and RO.4). Doppler radar antennas were mounted behind special plastic license plates on the front and rear of the vehicle (Figure RO.5) in the hope that this would make them inconspicuous to other drivers.

Figure RO.3. The main Data Acquisition System (DAS) unit mounted under the "package shelf" of the trunk.

Figure RO.4. The 100-Car Study DAS main unit shown without the top and front covers.

Figure RO.5. Doppler radar antenna mounted on the front of a vehicle, covered by one of the plastic license plates used for this study.

Other major components were mounted above and in front of the center rear-view mirror (Figures RO.6 and RO.7). These included an "incident" pushbutton that the subject could press whenever an unusual driving event occurred. An unobtrusive miniature camera for the driver face view was also contained in the housing for the pushbutton. The camera was invisible to the driver since it was mounted behind a "smoked" Plexiglas cover. The forward-view camera and the glare sensor were mounted behind the center mirror (Figure RO.7). This location was selected because it was unobtrusive and did not occlude the driver's normal field of view.

Figure RO.6. The incident push button box mounted above the rearview mirror. The portion on the right contains the driver face/left road view camera hidden by a smoked Plexiglas cover.

Figure RO.7. The mounting for the glare sensor behind the rearview mirror. Note the forward view camera as part of the same mounting assembly.

Subjects

One-hundred drivers who commuted into or out of the Northern Virginia/Washington, DC, metropolitan area were recruited as primary drivers for this study. They could either have their private vehicles instrumented or receive an instrumented leased vehicle to drive for the duration of the study. Drivers were recruited with flyers and classified ads. Drivers under the age of 30 who did not drive a vehicle of an appropriate make and model were given a leased vehicle (22 vehicles), while drivers who drove the appropriate makes and models had their private vehicles instrumented (78 vehicles). For allowing their vehicle to be instrumented, these participants received $125 per month and a bonus at the end of the study. Leased-vehicle drivers received free use of the vehicle, including standard maintenance, and the same bonus at the end of the study.

A few drivers were replaced for various reasons (for example, a move from the study area or repeated crashes in leased vehicles), so a total of 109 primary drivers were included in the study. Other family members and friends occasionally drove the instrumented vehicles, so data was also collected on 132 secondary drivers.

One goal of this study was to record as many crash and near-crash events as possible; this was facilitated by selecting subjects with higher than average crash- or near-crash risk exposure. Exposure was manipulated through the selection of a larger sample of drivers below the age of 25 and by the selection higher mileage drivers. The age and gender distribution of the primary drivers is shown in Table RO.1, while the distribution of miles driven by the subjects during the study is shown in Table RO.2. Although the data may be somewhat biased compared to the national averages in each case, a reasonably representative distribution was felt to be attained.

Table RO.1. Driver age and gender distributions.

Age Bins N % of total	Gender		
	Female	Male	Grand Total
18-20	9 8.3%	7 6.4%	16 14.7%
21-24	11 10.1%	10 9.2%	21 19.3%
25-34	7 6.4%	12 11.0%	19 17.4%
35-44	4 3.7%	16 14.7%	20 18.3%
45-54	7 6.4%	13 11.9%	20 18.3%
55+	5 4.6%	8 7.3%	13 11.9%
Total N Total Percentage	43 39.4%	66 60.6%	109 100.0%

Table RO.2. Actual miles driven during the study.

Actual miles driven	Number of Drivers	Percentage of Drivers
0-9,000	29	26.6
9,001-12,000	22	20.2%
12,001-15,000	26	23.9%
15,001-18,000	11	10.1%
18,001-21,000	8	7.3%
More than 21,000	13	11.9%

The 100-Car Study data sample was collected at one site (i.e., Northern Virginia/Metro Washington, DC) due to the need to restrict the geography such that vehicles could be "chased" (as previously explained) for data download. This area represents primarily urban- and suburban driving conditions, often in moderate to heavy traffic. Thus, rural driving, as well as differing demographics within the United States, are not well represented. The Northern Virginia/ Metro Washington, DC, was chosen as the data collection site primarily because the urban driving environment would provide a higher crash risk than rural areas, and also because of its close proximity to Blacksburg, VA.

A goal of the recruitment process was to avoid extreme drivers in either direction (i.e., very safe or very unsafe). Self-reported traffic violation and crashes data are provided for each age group in Figures RO.8 and RO.9. These data indicates that a diverse distribution of drivers was obtained. Note, however, that for the number of years they have driven (2 to 8), younger drivers have a similar number of violations and crashes as drivers who self-reported for 10 years. This

observation that younger drivers have a higher violation and crash rate is observed in database analyses.

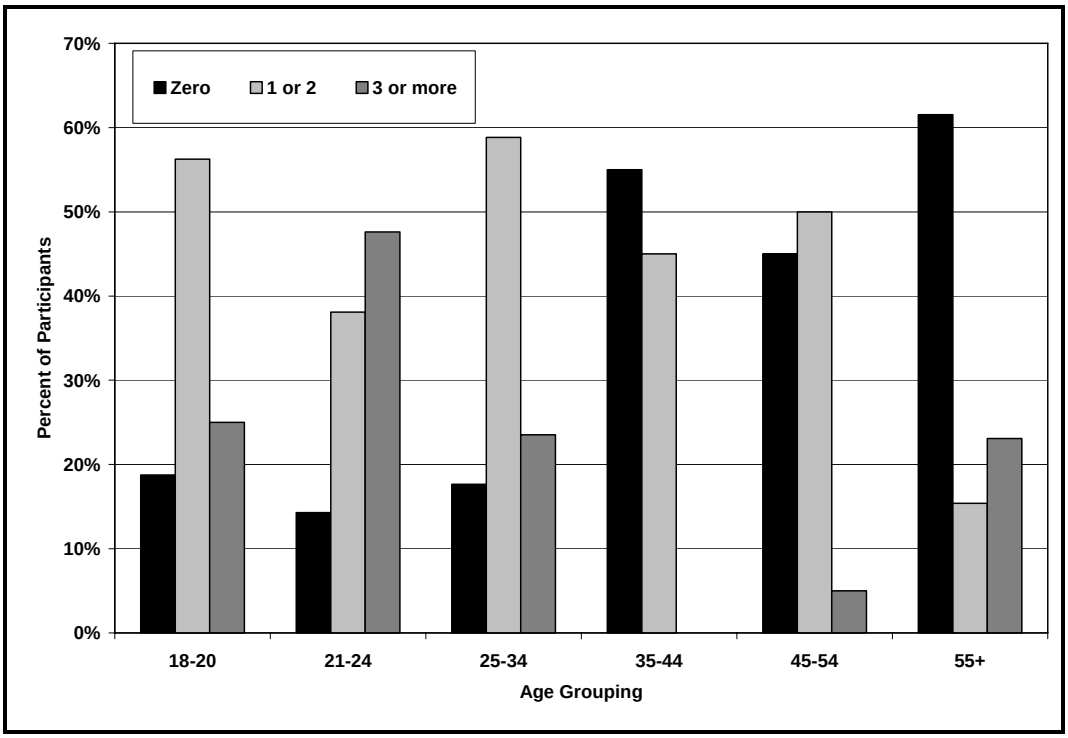


Figure RO.8. Number of self-reported traffic violations in the past 5 years as a percentage of driver age group.

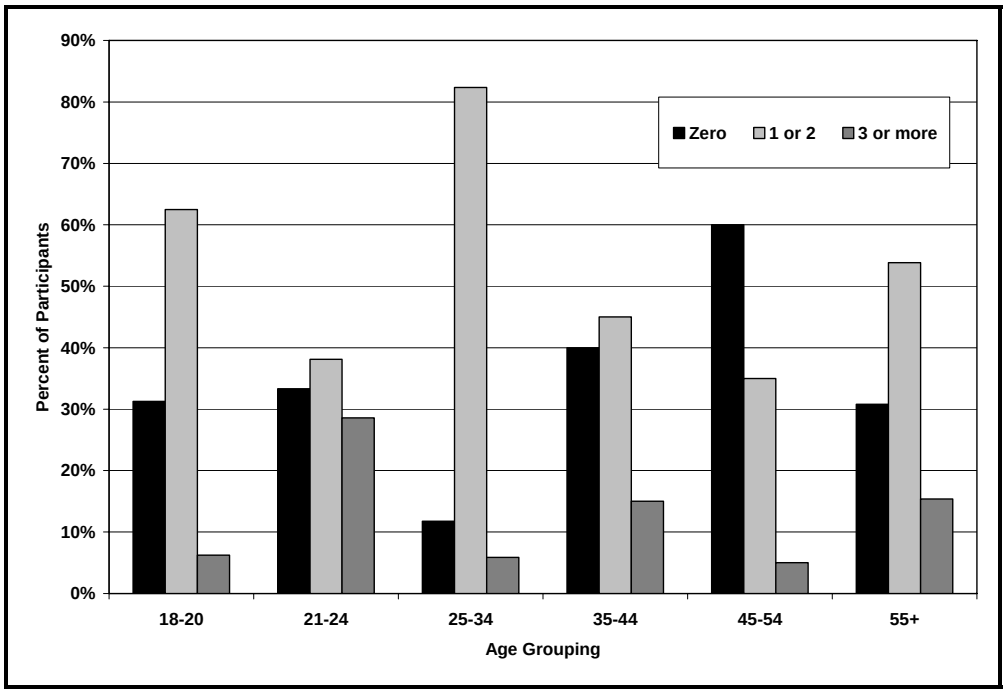


Figure RO.9. Number of self-reported traffic crashes in the past 10 years as a percentage of driver age group.

Vehicles

The number of vehicle types was limited for this study, since the complexity of the hardware required a number of custom mounting brackets to be manufactured. Six different vehicle models were selected based upon their prevalence in the Northern Virginia area. These included 5 sedan models (Chevrolet Malibu and Cavalier, Toyota Camry and Corolla, and Ford Taurus) and one SUV model (Ford Explorer). The model years were limited to those with common body types and accessible vehicle networks (generally 1995 to 2003). The distribution of these vehicle types is shown in Figure RO.10.

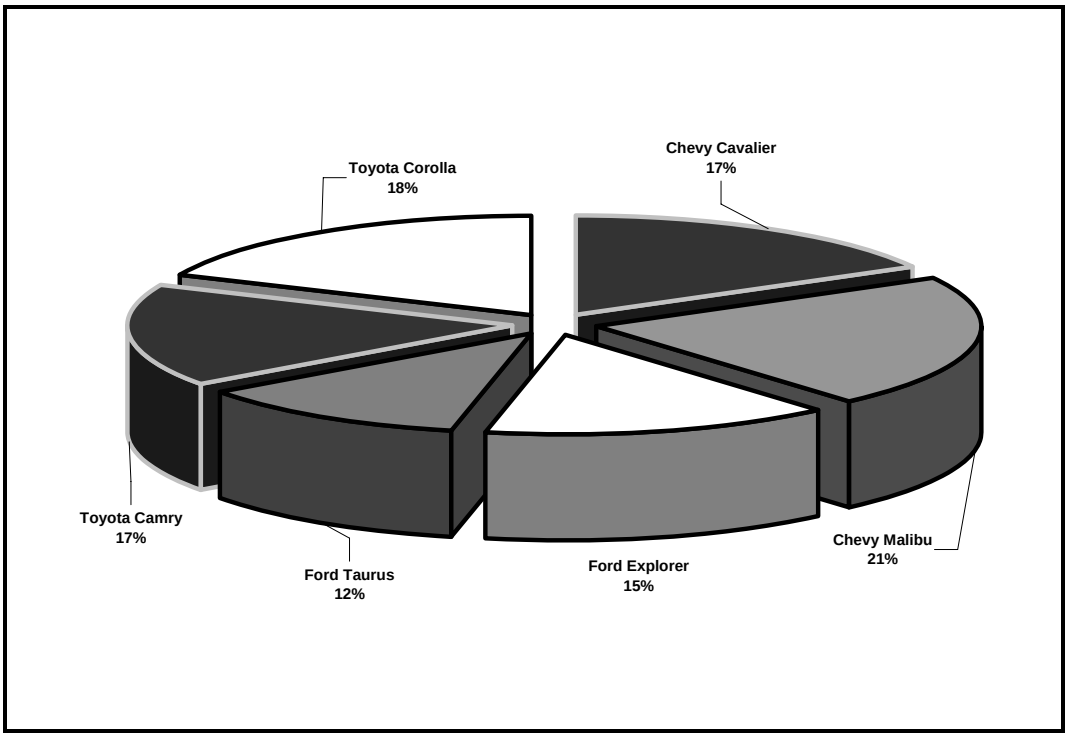


Figure RO.10. Distribution of vehicle makes and models driven during the study.

Results for the Initial 10 Project Goals

Ten specific goals were addressed as part of the initial data reduction and analysis of the 100-Car Study dataset. The results of each of the 10 analyses are summarized in the following sections.

GOAL 1: CHARACTERIZATION OF CRASHES, NEAR-CRASHES, AND INCIDENTS FOR THE 100-CAR STUDY

The purpose of the goal was to provide a "top-down" characterization of the reduced events. The events were characterized into three different levels of severity. Table RO.3 shows the relative frequency of each crash, near-crash, and incident for each conflict type (See *Glossary of Terms*). Of the 82 crashes, 13 either occurred during the initial computer 90-second initialization period, or contained incomplete data for other reasons (e.g., camera failure). There were a total of 69 crashes, 761 near-crashes, and 8,295 incidents for which data could be completely reduced. The first 8 conflict types shown in Table RO.3 accounted for all of the crashes, 87 percent of the

near-crashes, and 93 percent of the incidents. Therefore, these 8 conflict types were the focus of much of the *Goal 1* analysis.

Table RO.3. Number of crashes, near-crashes, and incidents for each conflict type.

Conflict Type	Crash	Near-crash	Incident
Single vehicle	24	48	191
Lead vehicle	15	380	5783
Following vehicle	12	70	766
Object/obstacle	9	6	394
Parked vehicle	4	5	83
Animal	2	10	56
Vehicle turning across subject vehicle path in opposite direction	2	27	79
Adjacent vehicle	1	115	342
Other	0	2	13
Oncoming traffic	0	27	184
Vehicle turning across subject vehicle path in same direction	0	3	10
Vehicle turning into subject vehicle path in same direction	0	28	90
Vehicle turning into subject vehicle path in opposite direction	0	0	1
Vehicle moving across subject vehicle path through intersection	0	27	158
Merging vehicle	0	6	18
Pedestrian	0	6	108
Pedalcyclist	0	0	16
Unknown	0	1	3

Unlike crash databases, all crashes are shown in Table RO.3, including non-police-reported, low-speed collisions. A "crash" was operationally defined for this study as "any measurable dissipation or transfer of energy due to the contact of the subject vehicle with another vehicle or object." One advantage of the naturalistic approach is that all of these events were recorded; however, it was necessary to develop crash severity categories in order to better understand the data. The 69 crashes were thus reviewed and placed into the following four levels:

- Level I: Police-reported air bag deployment and/or injury.
- Level II: Police-reported property damage only.
- Level III: Non-police-reported property damage only.
- Level IV: Non-police-reported low-g physical contact or tire strike (greater than 10 mph).

Therefore, the reader should keep crash severity in mind when reviewing this data. For example, 75 percent of the single-vehicle crashes were low-g-force physical contact or tire strikes. This type of crash, while indicative of loss of vehicular control, is not currently present in any crash database. This lack of representation is particularly important when considering the relationship between crashes, near-crashes, and incidents in Table RO.3. The breakdown of crash severity by crash type is shown in Table RO.4. As shown, the level I and II crashes provide a more

consistent ratio relative to near-crash events. These relationships were analyzed in much greater detail as part of other goals in this report.

Table RO.4. Crash type by crash severity category.

Conflict Type	Total	Level I	Level II	Level III	Level IV
Single vehicle	24	1	0	5	18
Lead vehicle	15	1	3	5	6
Following vehicle	12	2	2	5	3
Object/obstacle	9	0	1	3	5
Parked vehicle	4	0	0	2	2
Animal	2	0	0	0	2
Vehicle turning across subject vehicle path in opposite direction	2	1	1	0	0
Adjacent vehicle	1	0	0	1	0

The ability to detect crashes regardless of severity made it possible to examine the number of subjects who experienced a single crash versus the number who experienced multiple crashes during the 12- to 13-month data collection period. The number of crashes, near-crashes, and incidents experienced by the drivers is summarized in Table RO.5. As shown, 7.5 percent of drivers never experienced an event of any severity. In contrast, 7.4 percent of the drivers experienced many incidents and three or 4 crashes. As discussed in much greater detail as part of Chapter 4, *Goal 1*, a handful of subjects were very risky drivers and a handful of subjects were very safe drivers, reflecting a relatively normal distribution of events among drivers.

Table RO.5. Number and percentage of drivers involved in multiple events.

Number of Crashes	Percentage of Drivers	Number of Near-crashes	Percentage of Drivers	Number of Incidents	Percentage of Drivers
0	64.5%	0	16.8%	0	7.5%
1	21.5%	1	7.5%	1-5	9.3%
2	6.5%	2-4	27.1%	6-10	3.7%
3	3.7%	5-8	27.1%	11-15	0.9%
4	3.7%	9-12	3.7%	16-20	3.7%
More than 4	0.0%	13-24	13.1%	21-25	5.6%
		25-50	2.8%	26-30	4.7%
		More than 50	1.9%	31-40	8.4%
				41-50	7.5%
				51-100	16.8%
				101-150	16.8%
				151-200	11.2%
				More than 200	3.7%

Tree diagrams, like the one shown in Figure RO.11, were constructed to show the distribution of events. These diagrams outline the factors recorded for each conflict type. Additional diagrams for all conflict types are discussed in Chapter 4, *Goal 1*. Due to its size (over 400 pages), the full tree structure for the 100-Car Study events is shown in Appendix C.

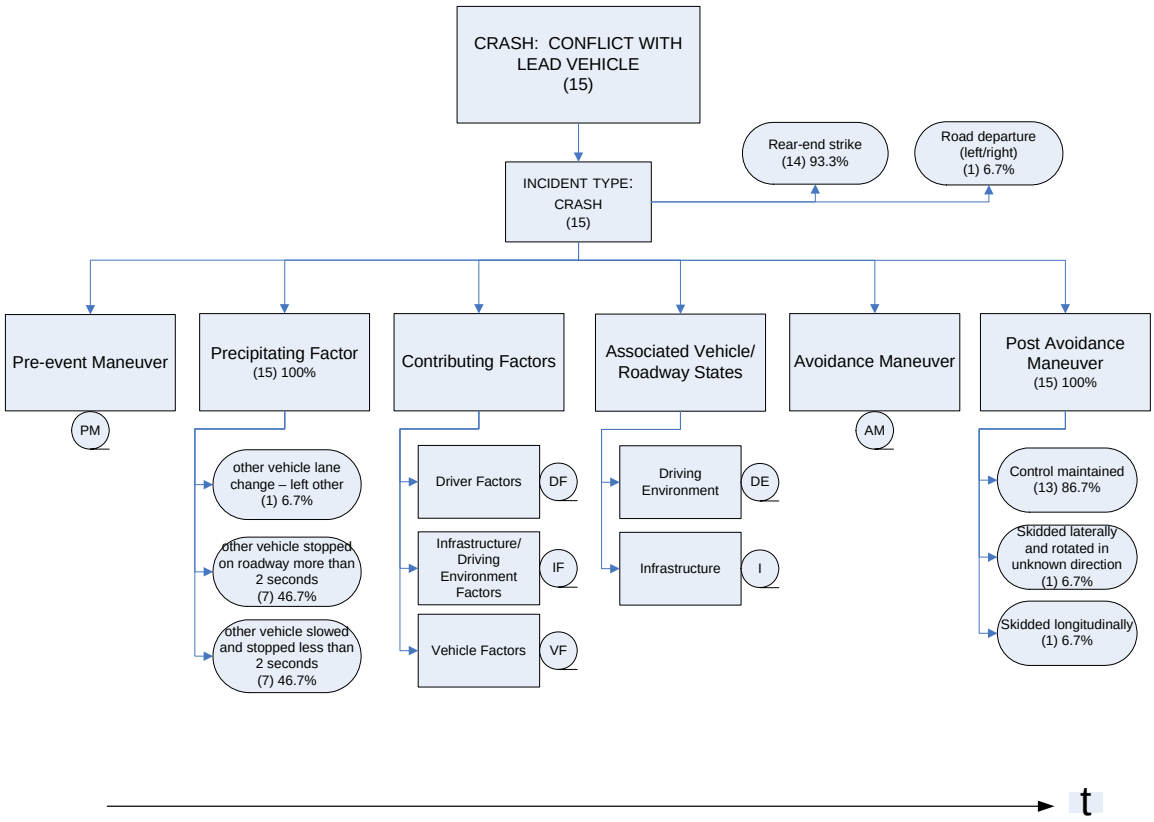


Figure RO.11. An example of a tree diagram used to delineate the contributing and associative factors for the 100-Car Study.

GOAL 2: QUANTIFICATION OF NEAR-CRASH EVENTS

The purpose of this research goal was to determine if near-crashes could be identified quantitatively using driving performance data. The 100-Car Study was the first study to capture a large set of near-crash events (there were over 800 near-crashes). Previous studies have operationally defined near-crashes based on subjective criteria to a large degree. This was also true for the current study. Near-crashes in the 100-Car Study were defined as a conflict event requiring a "rapid, evasive maneuver" in order to avoid a crash, although this definition was supplemented by quantitative guidelines to help data analysts decide when an event was a near-crash (e.g., a longitudinal deceleration of at least 0.5 *g*). However, many safety applications would greatly benefit from a reliable, purely quantitative definition of a near-crash event.

An analytical attempt was made to develop a near-crash criterion for several applications, including:

1) An in-vehicle trigger to collect data on near-crash events in a large scale study where continuous data collection would not be feasible.
2) An in-vehicle trigger to warn a driver of a severe conflict scenario (i.e., for a collision avoidance warning system).

These analyses were somewhat successful in developing a near-crash data-based trigger for a large-scale data collection effort, but were not successful in determining a crash warning boundary.

One reason for suboptimal near-crash boundary performance was simply noisy sensor data. In some cases, radar units missed the critical target because the target did not appear in the radar's field of view. This phenomenon commonly occurs during lane changes when the lead vehicle is lost as the subject vehicle turns into the new lane. A specific example of this occurred during the telephone pole crash where the driver swerved to the right to miss the lead vehicle and hit a telephone pole instead. As the subject vehicle grazed the rear corner of the lead vehicle, the lead vehicle left the radar unit's field of view. Alternatively, radar units detected non-critical targets, such as guard rails, when the road geometry was off-angle. For these cases, the current level of false alarms and misses might be reduced with more sophisticated technology and algorithms.

Nevertheless, the data clearly showed that development of purely quantitative near-crash criteria is not currently feasible for most cases. One major reason is that the kinematic signatures associated with near-crash events are virtually identical to many common driving situations that are *not* indicative of crash risk. An example of this is shown in Figure RO.12. Shown is a range/range rate plot that includes three boundary types: the two green lines are approximations of graded warning and advisory boundaries used in recent research (Kiefer et al., 2003). The black line is a minimum error boundary that could be used to automatically detect near-crash events in a large scale study. As shown, there are many invalid cases (red dots indicating no conflict present), particularly above the minimum error boundary of the range/range rate plot.

The implication of this analysis for large-scale naturalistic data collection is that video data verification of dynamically triggered events will likely be necessary, at least for the foreseeable future. However, as discussed in the report *Goal 10: Evaluation of the Performance of the 100-Car Naturalistic Driving Study Data Reduction Plan, Triggering Methods, and Data Analysis* (separate report), such verification is neither difficult nor expensive relative to the overall

collection effort of such large-scale field tests, given current video technology. From a large-scale naturalistic study perspective, crash detection is reasonably straightforward since there is often a greater than 1.0 *g* peak deceleration when a crash occurs. The detection of near-crash cases is more problematic. However, depending on the size of the study, it may be reasonable to make an *a priori* decision to capture in the range of 70 percent of 25,000 or 30,000 near-crash events if the false alarm rate can be reduced to around 10 percent. Even with a higher false alarm rate, the cost of each false alarm would be fairly low given data reduction tools similar to those used in this study. For the current study, a trained reductionist was able to distinguish between valid and invalid conflicts at the rate of about 50 per hour using video data. This topic is further discussed in the *Goal 10 Report (separate report).*

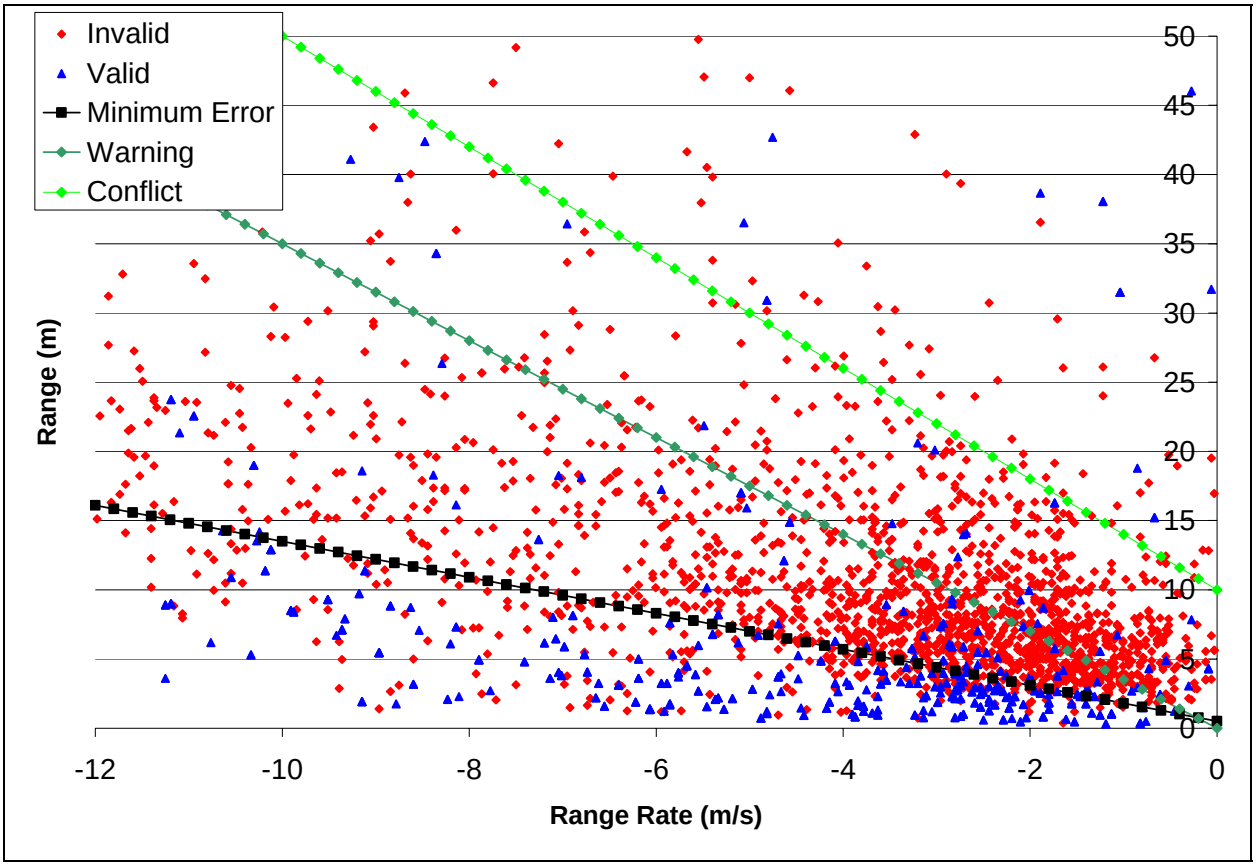


Figure RO.12. Point of greatest threat with lead vehicle for all crashes and near-crashes, and a random sample of invalid events. The boundaries shown are approximations of warning and conflict boundaries used as part of a forward collision warning algorithm (Kiefer et al., 2003) and a minimum error boundary calculated for this dataset.

GOAL 3: CHARACTERIZATION OF DRIVER INATTENTION

Historically, driver distraction has typically been associated with secondary tasks such as dialing a cell phone, conversing with a passenger, and adjusting the radio. Driver distraction has been said to lead to driver inattention. *Drowsiness* has been described as another cause of driver inattention. With the video data available in this study, new categories of "*driver inattention*" were discovered. The two new categories were "*driving-related inattention to the forward*

roadway" and "*nonspecific eyeglance.*" "*Driving-related inattention to the forward roadway*" involves the driver checking the speedometer, rear-view mirrors, or blind spots. This new category was added after viewing numerous events for which the driver was clearly paying attention to the driving task, but was not paying attention to the *critical aspect* of the driving task (i.e., the forward roadway) at an inopportune moment.

Further eyeglance analysis was performed manually by data reductionists using only crashes and near-crashes in the 100-Car Study database. The "*nonspecific eyeglance away from the forward roadway*" describes cases for which the driver briefly glances away from the roadway, but at no discernable object or person. For this project, eyeglance reduction was performed for crash and near-crash events only, so this category can only be used for the more severe events. The four inattention categories combined (secondary task, drowsiness, inattention to forward roadway, and nonspecific eyeglance) suggest that driver's glances away from the forward roadway may contribute to a much greater percentage of events than has been found in previous studies (Campbell, Smith, and Najm, 2003). As shown in Figure RO.13, 78 percent of the crashes and 65 percent of the near-crashes had one of these four inattention categories as a contributing factor.

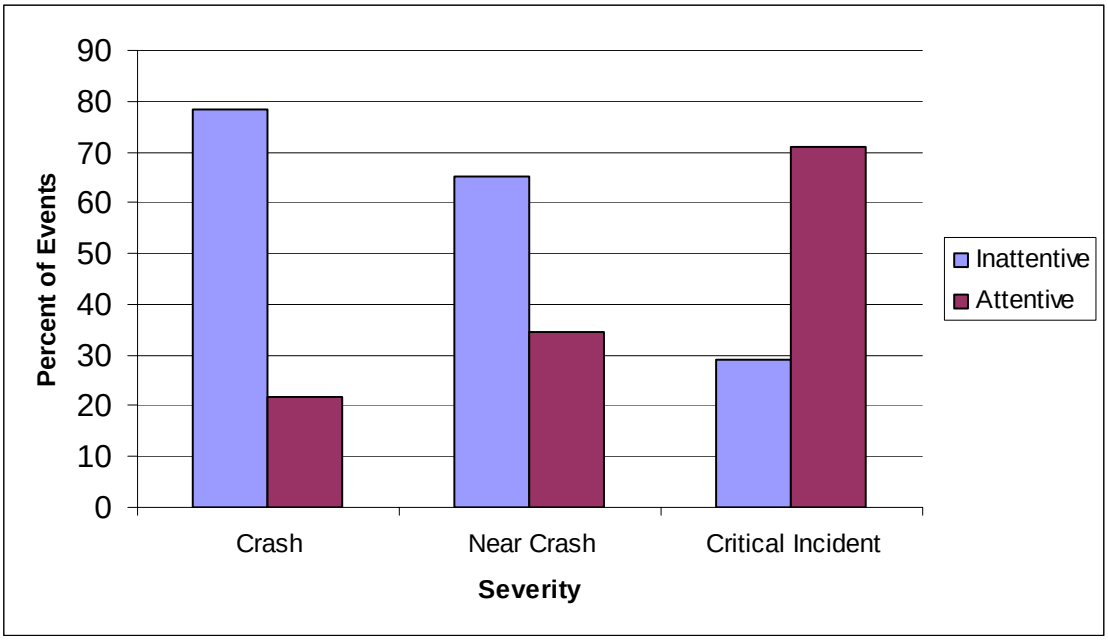


Figure RO.13. Percentage of events for attention by severity level.

An analysis of these types of inattention revealed that secondary task distraction was the largest of the four categories (see Appendix D for more complete descriptions of the inattention categories). The secondary tasks that generally contributed to the highest percentages of events (Figure RO.14 for crashes and near-crashes) were wireless devices (primarily cell phones), internal distractions, and passenger-related secondary tasks (primarily conversations). It is important to note that exposure is not considered in these data. An analysis of frequency of device use is currently being conducted for a future report that will quantify exposure-based risk.

Figure RO.15 shows a breakdown of the wireless device tasks (see Appendix D for more thorough descriptions of cell phone categories). All of the crashes and a majority of the near-

crashes and incidents associated with wireless devices occurred during a cell phone conversation, although the dialing task was also relatively high in term of total conflicts. Although these data do demonstrate factors that contribute to these wireless task events, there is still a need for exposure data to adequately assess the risk associated with these wireless device tasks.

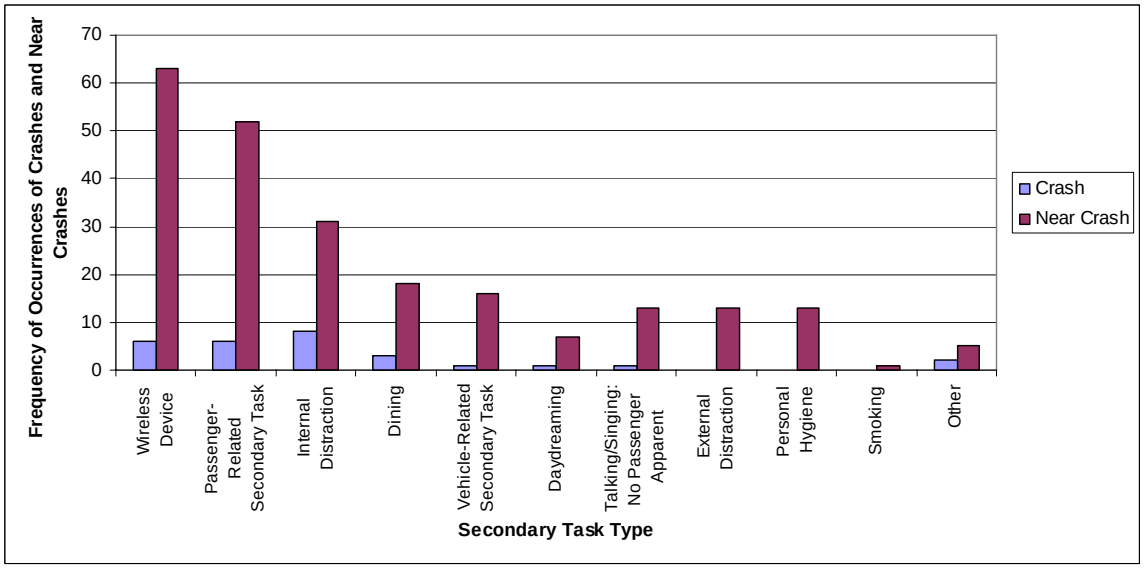


Figure RO.14. Comparison of crashes and near-crashes the frequency of occurrences of the presence of a distracting agent as a contributing factor.

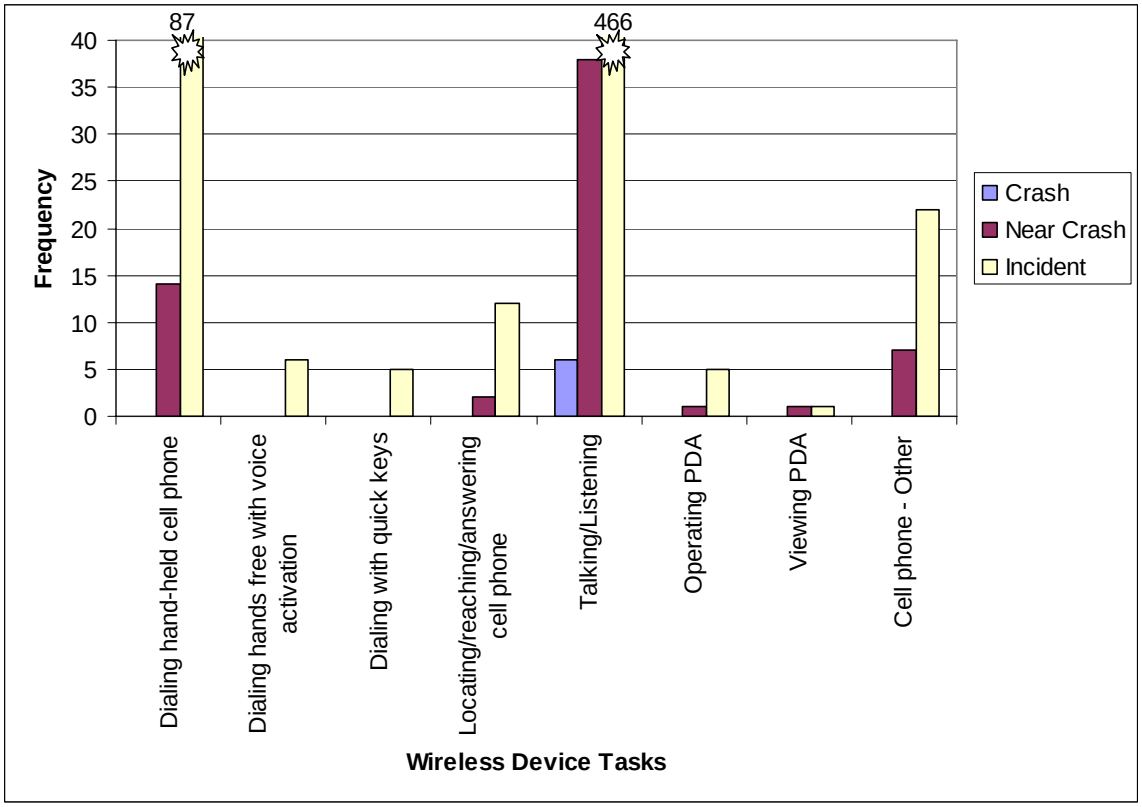


Figure RO.15. Frequency of events for which the contributing factor was wireless device use.

GOAL 4: DRIVER BEHAVIOR OVER TIME

The questions addressed in this goal were intended to explore issues of whether driver behavior in an instrumented vehicle changed over time. The units of time used were weeks (weeks 1 through 50) and hours (the first 50 hours). The issues explored were: (1) driver behavior in a newly instrumented leased vehicle in the first weeks as compared to the last few weeks of the study; (2) driver behavior in the first few hours of driving; and (3) driver behavior for the same driver in four weeks of leased vehicle driving versus four weeks of private vehicle driving. The relative risk (RR) analysis technique was borrowed from the field of epidemiology, and required that there be both an exposed and unexposed condition and a comparison and baseline time period. For these questions, the exposed condition was the leased vehicle and the unexposed condition was the private vehicle, since the private vehicle drivers kept driving their usual vehicles while the leased vehicle drivers were exposed to a new vehicle.

For the 100-Car Study dataset, one of the primary questions in calculating RR was the decision regarding which period or periods of time to use as a control period for each question of interest. A preliminary examination of the frequency data indicated the presence of random fluctuations in the frequency of these events for any given day or month. These random fluctuations are a result of the relatively infrequent occurrence of crashes, near crashes, and incidents. To control for these random fluctuations, a decision was made to use an average of the final time periods for

each question of interest as the control time period. The baseline time periods were an average of weeks 41-50 for the yearly comparison, hours 41-50 for the hourly comparison, and weeks 2-4 for the leased versus private vehicle comparisons.

There was a potential confound between leased vehicles and driver age, in that none of the leased vehicle drivers were over the age of 30, so any results from the leased vehicle analyses may have been confounded by age. An age analysis confirmed that younger drivers did indeed have an elevated RR as compared to older drivers, but that the effect was not as large as the effect for leased versus private vehicles. However, there were approximately 25 younger drivers of both leased and private vehicles. The age distributions of these two groups were quite similar, so this age-matched set of drivers was used for questions relating to weekly data. A similar matched set of younger drivers was used for hourly data, while the vehicle adaptation questions used a matched set of switch drivers (those who moved from a private vehicle to a leased vehicle at the end of the study).

Driving Behavior Over the Course of a Year

The issue of interest here was the driver adaptation process for leased vehicles and privately-owned vehicles with instrumentation over the course of the study. It was expected that drivers would be most adapted to their vehicles and to the instrumentation by the end of the study, so weeks 41-50 were used as the baseline time period. Adaptation to the vehicle instrumentation was explored in terms of the both the average number of events per vehicle and the number of events per mile for private and leased vehicles.

When the number of events were examined, it became obvious that although there was not any appreciable change in the number of events over the course of the year, there was a consistently higher risk for leased vehicle drivers as compared to private vehicle drivers (Figure RO.16). The calculated RRs for this graph were above 1 for every time period, and sometimes above 1.5, but the 95th percentage lower CI of the RR is below 1, so these differences are likely not significant. However, the trends seen for every analysis performed for Goal 4 were noticeable in magnitude, were consistent over time, and always showed the same effect (leased vehicle higher than private vehicle).

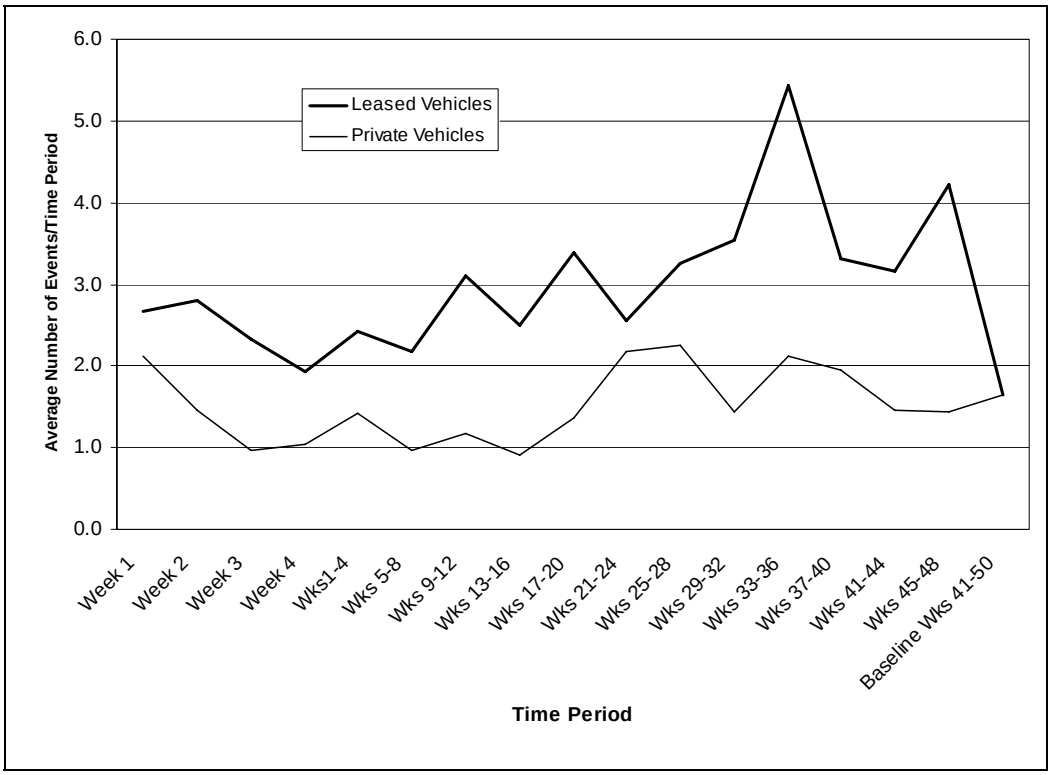


Figure RO.16. Mean number of events for a matched set of younger drivers for leased and private vehicles over weeks 1-50 of the study.

It was hypothesized that leased vehicle drivers may be willing to take risks leading to near-crashes and incidents with these vehicles, since they were not responsible for insurance or repairs, and had no ownership interest in the vehicles. The crash data supported this hypothesis in that the RR was lower for crashes than near-crashes and incidents. Based on these results, one might expect that if one were to transfer the leased vehicle drivers into their own private vehicles in which they would be responsible for repairs, insurance, etc., that their event levels would drop to the same levels shown for private vehicle drivers.

Driving Behavior Over the First 50 Hours

The next questions were designed to determine whether drivers experienced an increase in valid events over the first few hours of driving a newly instrumented vehicle. It was hypothesized that the drivers would drive more carefully and experience fewer events when they were aware of the cameras, and that they would revert to normal behavior as time went on. If a point in time can be identified at which drivers adapted and began acting more naturally, this would be useful information for future instrumented vehicle studies of naturalistic driving. Previous experience at VTTI has indicated that drivers adapt amazingly quickly to the instrumented vehicle (perhaps within minutes, even in an unfamiliar vehicle), but the question has never been empirically analyzed as was attempted here. A matched set of younger drivers was also used for these questions.

As before, even when controlling for age to the degree possible, leased vehicles experienced a greater mean number of events for nearly every time period studied. The only exceptions were hours 1 and 4, in which the leased and private vehicles experienced nearly identical mean

numbers of events. It did appear that drivers of both vehicle types were being very careful during the first hour with a newly instrumented vehicle. These results provided support for the thesis that drivers are more careful when first using an instrumented vehicle, although the effect appears to wear off after the first hour. The dataset did not provide a breakdown by minutes, so it was not possible to tell whether this occurred within the first 5 minutes, the first half hour, or at the end of the first hour.

Performance for Same Driver for Four Weeks in Private and Leased Vehicles

The purpose of these questions was to investigate the driver adaptation process to an unfamiliar vehicle for the same driver in a leased vehicle versus a privately-owned vehicle (both instrumented). Only switch drivers for whom matched data were available for each week were used, resulting in a perfectly matched set of drivers for each week.

The data did not indicate any clear trend of adaptation to a new vehicle. When examining the leased versus private vehicle question, however, the analyses using perfectly matched sets of switch drivers had similar results to the previous analyses. Even when the same driver was switched from a private vehicle to a leased vehicle, there were still a greater number of events in the leased vehicle than in the private vehicle. As shown in Figure RO.17, the same younger drivers had a consistently higher mean number of events over weeks 1-4 when driving the leased vehicle as compared to weeks 1-4 in their private vehicle.

If the increased number of events in leased vehicle driving for the same driver was due to vehicle unfamiliarity, this effect was not extinguished over the first four weeks. Based on the yearly results, the higher numbers for leased vehicles likely had very little to do with adaptation since after 50 weeks there were still more events for leased vehicles as compared to private vehicles. The results of the per mile analysis were in close agreement with the per vehicle analysis.

In order to further explore the issue of adaptation to a new vehicle, switch driver data were also examined over the first 10 hours of driving. Six of the switch drivers had data for each of the first 10 hours. There were no events in the first hour of driving for either leased or private vehicle driving for these 6 drivers, providing support for drivers being more careful during the first hour after beginning to drive an instrumented vehicle, but not providing real support for adaptation to a new vehicle. Altogether, these 6 drivers experienced 9 events in the first 10 hours of driving their own private vehicle and 18 events in the first 10 hours after switching over to a leased vehicle. Six of the 9 private vehicle events and all 18 of the leased vehicle events involved younger drivers, providing evidence that individual younger drivers may have more trouble adapting to a new vehicle.

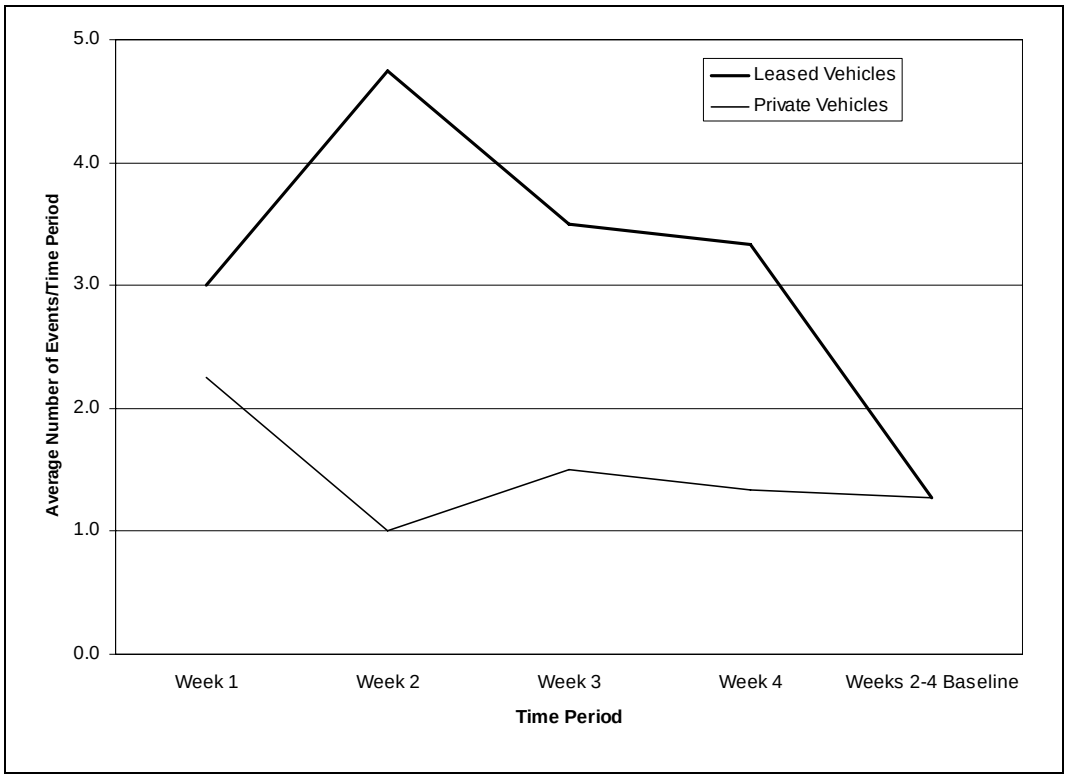


Figure RO.17. Matched set of younger switch drivers: leased versus private vehicle mean number of events for weeks 1-4.

GOAL 5: REAR-END CONFLICT CAUSAL FACTORS AND DYNAMIC CONDITIONS

A primary purpose of this research was to examine the contributing and associative factors for rear-end events (specifically, *conflict with lead vehicle* and *conflict with following-vehicle* events). Recall that the 100-Car Study instrumentation had both forward and rear facing radars and cameras, allowing analysis of both types of data. Nonetheless, much more data were available for "*conflict with lead vehicle*" cases, since other sensors and cameras were available in the *subject vehicle* relative to the *following vehicle.*

The frequency of lead vehicle and following-vehicle events by level of severity was determined for the driver data included in the analyses. For the lead vehicle conflict case, the resulting dataset contained 13 crashes, 268 near-crashes, and 4,747 incidents. For the following-vehicle conflict case, the resulting dataset contained 9 crashes, 30 near-crashes, and 239 incidents.

The four questions answered for this goal addressed driver characteristics, kinematic characteristics, contributing factors, and corrective action for RE events. All data were presented in the form of event rate per million vehicle miles traveled (MVMT). The 5 RE scenarios considered were *LV accelerating*, *LV moving at slower constant speed*, *LV decelerating*, *LV stopped less than or equal to 2 seconds,* and *LV stopped greater than 2 seconds.* A summary of some of the most important findings is outlined here. The full set of results is shown in Chapter 9, *Goal 5*.

The frequency of lead-vehicle events for each of 5 lead-vehicle scenarios is shown in Table RO.6. It can be seen that the most common scenario for incidents was *LV decelerating*, followed by *LV stopped greater than 2 seconds.* For near-crashes, the most common scenario was again *LV decelerating*, followed this time by *LV stopped less than or equal to 2 seconds.* It is noteworthy that although *LV decelerating* was the most common scenario for incidents and near-crashes, there were no crashes for this scenario. All of the crashes occurred in circumstances for which the LV was stopped when the crash occurred, either more than 2 seconds (6 crashes) or 2 seconds or less (7 crashes). There were a fairly small number of incidents and near-crashes for LVs moving at a slower, constant speed. There were only 8 incidents for LV accelerating, and no crashes or near-crashes for this scenario.

Table RO.6. Frequencies for the 5 RE lead-vehicle scenarios by event severity.

Severity	LV accelerating	LV moving slower, constant speed	LV decelerating	LV stopped ≤ 2 s	LV Stopped > 2 s
Incident	8	119	2,436	989	1,195
Near-Crash	0	5	148	74	41
Crash	0	0	0	**7**	**6**

Following Vehicle Data

As was true for the lead-vehicle scenarios, the following-vehicle events were concentrated in the *SV decelerating* scenario. The next most common scenarios of *SV stopped less than or equal to 2 seconds* and *SV stopped greater than 2 seconds* were nearly equal in terms of frequency. Recall that for this scenario, the 100-Car Study subject vehicle (SV) was considered to be the lead vehicle and was struck from behind by a following vehicle. Table RO.7 presents the overall number of following-vehicle events. Note that in this case, unlike the lead-vehicle case, the *subject vehicle* was still decelerating at the time of a collision in 4 of the 10 crash cases. In the other 6, like the lead-vehicle case, the *subject vehicle* was stationary.

Table RO.7. Frequencies for the 5 RE following-vehicle scenarios by event severity.

Severity	SV accelerating	SV moving slower, constant speed	SV decelerating	SV stopped ≤ 2 s	SV Stopped > 2 s
Incident	1	21	207	48	63
Near-Crash	1	0	26	15	0
Crash	0	0	**4**	**2**	**4**

Driver characteristics were examined to assess the role of age and gender for lead-vehicle and following-vehicle events. The only distinct trend in the lead-vehicle age data was that 18-20-year-olds had the highest rate of incidents and near-crashes per mile for each of the 5 scenarios.

The next analyses considered the kinematic conditions for RE lead-vehicle events. The kinematic data at the onset of the precipitating factor was used for these analyses. For lead-vehicle events, incidents and near-crashes had the highest rates for moderate speeds during the

event epoch of 21-40 mph, while crashes had the highest rates at lower speeds of 0-20 mph. The high incident and near-crash rates for the moderate speed ranges likely reflect the prevailing speed limits and high traffic density present in the northern Virginia area where the study was conducted.

For the most common following-vehicle scenario of *SV decelerating*, the speed ranges of 11-20 and 21-30 mph had the highest rates of incidents. This is a somewhat lower speed range than was found for the lead-vehicle incidents. For near-crashes, the speed ranges for the *LV decelerating scenario* with the highest rates were 21-30 and 31-40 mph, which may be an indicator that increasing event onset speed results in increased event severity

The environmental and roadway contributing factors for lead-vehicle RE events was considered next. *Traffic density* was related to the highest incident rate by far of any of the environmental and roadway contributing factors for all 5 RE lead-vehicle scenarios. *Relation to junction* had the next highest rate for all 5 scenarios, followed by *traffic control*, *light*, and then *weather*. The relative rank of rates within each scenario was very consistent.

Corrective actions for lead-vehicle events were considered next. For the lead-vehicle *LV decelerating* and *stopped* scenarios, braking (no lockup) dominated the rate data by factors of around 10 to 1. The next highest rates were for braked and steered to right, braked and steered to left, and braking (lockup unknown). When the LV in a RE event was stopped, the SV response overwhelmingly involved some sort of braking activity, usually without steering. For *LV decelerating*, steering left and steering right also had fairly high rates, although the overwhelming choice was still braking. For *LV moving at slower constant speed*, a quite different kinematic situation, braking (no lockup) still had the highest rate, but it was nearly equaled by braked and steered to right and no avoidance maneuver.

Additional insight into RE events can be found in Chapter 10, *Goal 6*, and Chapter 11, *Goal 7*. The relationships between the relative frequency of crashes, near-crashes, and incidents for RE events are explored using Heinrich's Triangles in Chapter 12, *Goal 8*.

GOAL 6: LANE CHANGE CAUSAL FACTORS AND DYNAMIC CONDITIONS

As stated in *Goal* 5, a primary goal for the 100-Car Study was to determine the causes and contributing factors associated with RE crashes. Understanding the pre-event maneuvers and precipitating factors that, in conjunction with other contributing factors, lead to RE crashes is important for fully understanding the rear-end crash problem. The purpose of the analyses for Chapter 10, *Goal 6* was to understand the degree to which lane change events, such as cut-ins, lead to rear-end conflicts. This has important implications for the design of future forward collision warning systems, since a cut-in vehicle may not provide a radar signature until very late in a conflict scenario. To begin to understand this issue, the RE conflict data were analyzed for both lead-vehicle (i.e., subject vehicle as following vehicle) and following-vehicle (i.e., subject vehicle as lead vehicle) scenarios. Frequency distributions were generated to identify the rate that these types of scenarios occurred per MVMT, the initial kinematic conditions that occurred for each, and the contributing factors that played a role for each type of scenario.

No crashes occurred when there was a lane change as a precipitating factor in front of the subject vehicle or when there was a lane change behind the subject vehicle. There were, however, 64

near-crashes and 324 incidents that occurred when there was a cut-in to the lane in front of the subject vehicle as compared to only 4 near-crashes and 77 incidents for which the SV changed behind a lead vehicle. As will be described in a later chapter of this report, the subject vehicle drivers were judged to more often be impaired (30 incidents more), distracted (44 incidents more) and make proficiency-related errors (e.g., inappropriate reaction; 55 more) for the near-crashes and incidents in which they were cut-off than cut-off events in which they were attentive. This seems to support the finding that at least two elements are required for a conflict to occur; a precipitating factor plus another contributing factor (often driver state-related). In this case, there were fewer events when the subject vehicle was the cut-in vehicle because the drivers were presumably more alert and attentive when they were actively performing the lane change maneuver.

The lane-change-related SV striking events were analyzed according to age group. For lane change incidents, the 18-to-20-year-olds had the highest rate for *LV lane change in front of SV*, while the other age categories had fairly equal rates for this scenario. For the *SV lane change behind LV* scenario, the 21-to-24-year-olds had the highest rate, although the rates were fairly even across age groups. For near-crashes (Figure RO.18), 18-to-20-year-olds again had the highest rate for the *LV lane change in front of SV* scenario by a factor of nearly 2 to 1 over the next highest age group (35-to-44 year olds). Rate data for SV striking events by both age and gender were considered next. The only clear pattern that emerged was for the 45+ age group. The male drivers in this age group had a rate that was more than twice as high as that for female drivers for both *SV lane change behind LV* and *LV lane change in front of SV*. The gender comparisons were fairly even across the other age groups. Almost all of the near-crashes were of the *LV lane change in front of SV* type. Males had a noticeably higher near-crash rate than females for this scenario in the 18-to-24 and 45+ age groups, while females had a higher rate in the 25-to-44 age group.

For lane-change-related SV struck events, three age categories (18-to-20, 25-to-34, and 35-to-44) had incident rates that were 1.5 to 3.5 times as high as the other three age groups (Figure RO.19). For near-crashes, the three younger age groups had rates 1.5 to 4 times as high as the three older age groups. When both age and gender were considered, females had at least twice the rate as males for each of the three age groups. A similar pattern was observed for near-crashes, except that the 18-to-24-year-old males and females had virtually identical rates and the remaining differences for the other age groups were by at least a 4 to 1 margin (females higher than males).

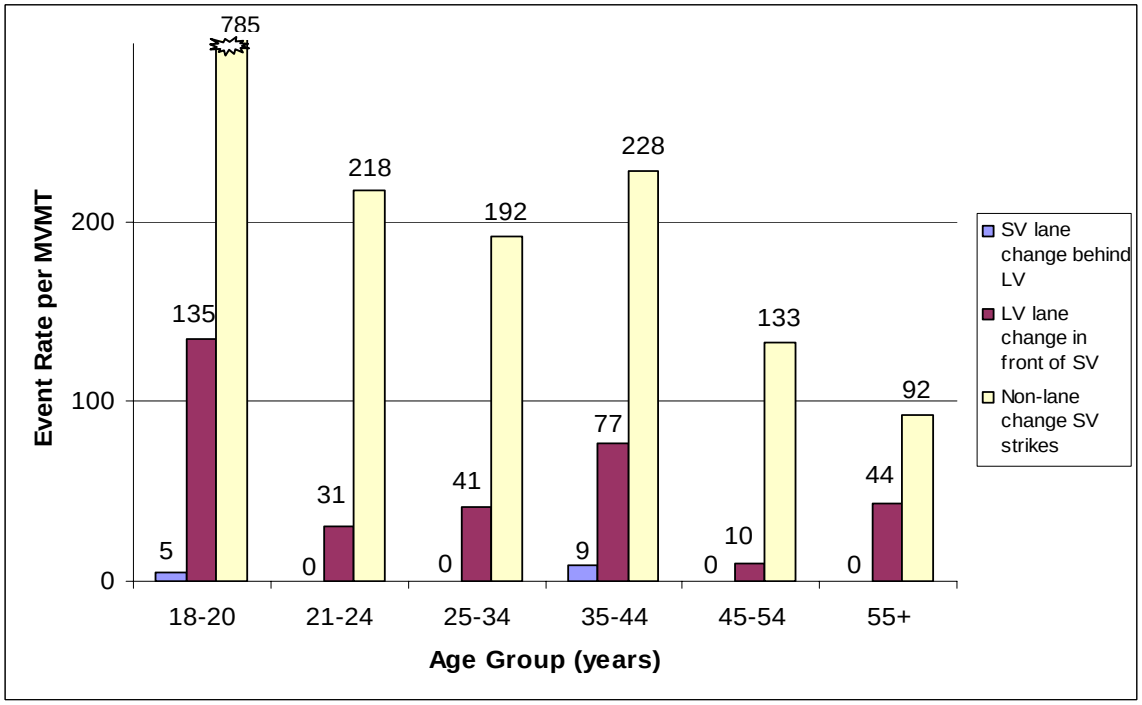


Figure RO.18. Rate per MVMT for SV striking by age and lane change maneuver for near-crashes.

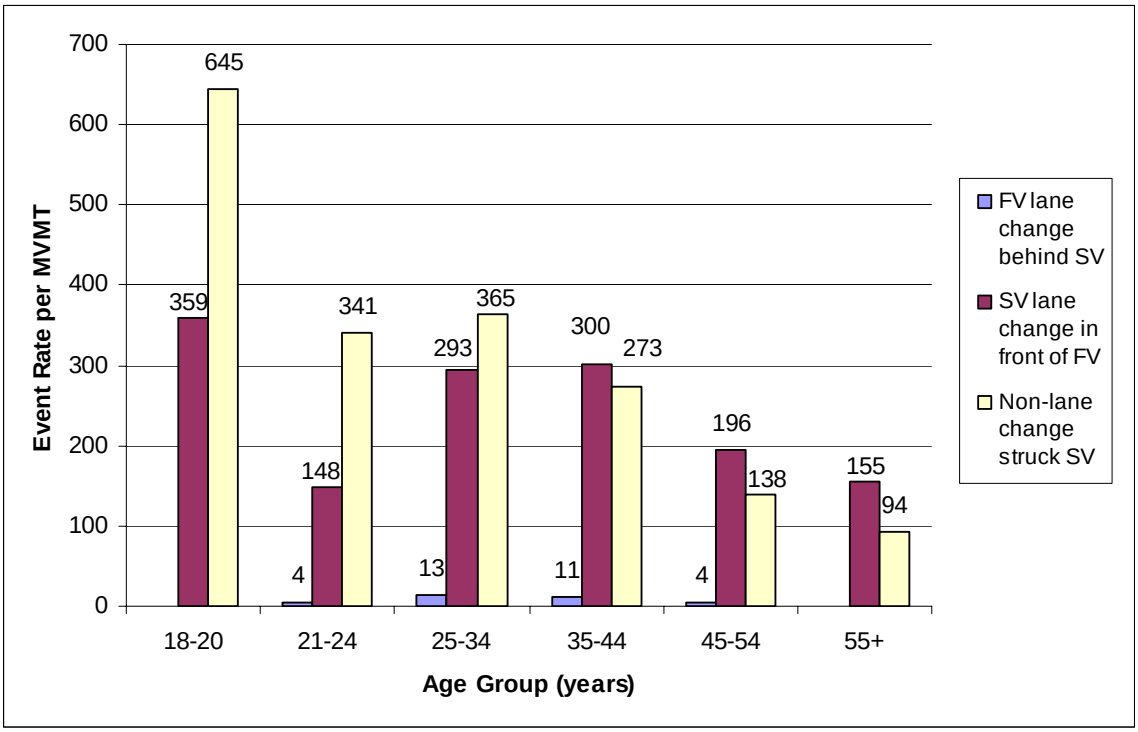


Figure RO.19. Rate per MVMT for struck SV by age and lane change maneuver for incidents.

Roadway and infrastructure factors were considered next. Traffic density was the factor with the highest incident rate for all four lane-change-related scenarios, by a factor of at least four in all but one case. Going back to the frequency data, 90 percent of all lane-change-related incidents were coded with traffic density as a contributing factor. Light, traffic control, and relation to junction were second, third, or fourth most important for all four scenarios. For the non-lane-related incidents, the highest rates were observed for traffic density, relation to junction, traffic control, and then light, quite a different pattern than was seen for the lane-change-related incidents. When the near-crash data were examined, the lane-change-related scenarios followed similar patterns as for incidents. For near-crashes, the non-lane-related rate pattern was more closely aligned to the lane-change-related near-crash data than to the non-lane-related incident data.

Driver contributing factors were also examined. Driver proficiency (as defined in the glossary) showed up as a prominent factor for the *SV struck* incident scenarios (higher than the next highest driver factor rate by 4 to 1). For the *SV striking* scenarios, the incident rates for driver factors were fairly even within each scenario, and driver proficiency was not even the top factor for the *SV lane change behind LV* scenario (the highest rated factor was willful behavior). When the near-crash rates were examined, driver proficiency and driver distraction also had the highest rates.

GOAL 7: INATTENTION FOR REAR END LEAD-VEHICLE SCENARIOS

The prevalence of distraction was of particular interest in the analyses of rear-end conflict contributing factor. The degree to which an unalerted driver can be warned and make a proper response is an important factor in developing rear-end crash countermeasures. The 100-Car Study data can provide great insight into the degree to which distraction is an issue in such conflicts. The important finding in this regard is that 93 percent of all lead-vehicle crashes (14 out of 15) involved *inattention to the forward roadway* as a contributing factor (Figure RO.20). Note also that a majority of the near-crashes have inattention listed as a contributing factor. Approximately one-third of the incidents have inattention listed as a contributing factor. The effect is nearly perfectly linear, and seems to indicate a strong correlation between inattention and increased severity for lead-vehicle rear-end events.

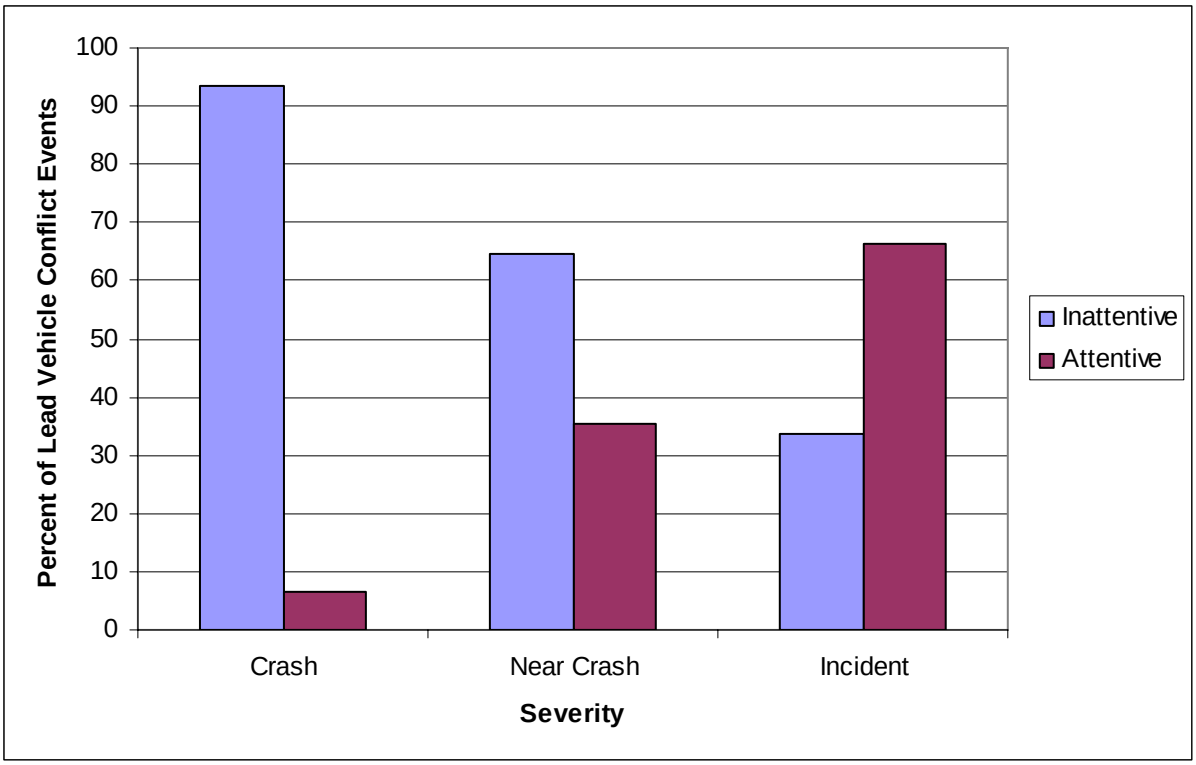


Figure RO.20. Percentage of lead-vehicle events for which inattention was listed as a contributing factor (includes the nonspecific eyeglance events for crashes and near-crashes).

Figures RO.21 and RO.22 shows the breakdown of the lead-vehicle conflict kinematic scenarios for the crash and near-crash events, respectively. As shown in figure RO.21, in 13 of the 14 *conflict with lead-vehicle crashes* the driver was inattentive, and all 14 crashes the lead vehicle was stopped when struck. For the near-crash events (Figure RO.22), the majority of the drivers were inattentive, but the largest lead-vehicle kinematic category was *lead vehicle decelerating.*

Taken together these results indicate that drivers have sufficient awareness and ability to perform evasive maneuvers when closing rates are lower and/or expectancies about the flow of traffic are not violated.

Figure RO.22 shows the frequency of each source of inattention for all secondary task categories. This allows comparison of the actual contribution of each of these sources of inattention to lead-vehicle conflicts. Wireless devices (primarily cell phones, but including a few PDA events) were the most frequent contributing factor for lead-vehicle events, followed by passenger-related inattention. The trend was very similar for near-crashes. Interior distractions were the most frequent source of inattention for crashes. Cell phone use was a dramatically more frequent contributor to incidents and near-crashes than any other secondary task, but did not contribute to any lead-vehicle conflict crashes. Nonetheless, cell phone use did contribute to several crashes of other types, as reported in other chapters of this report.

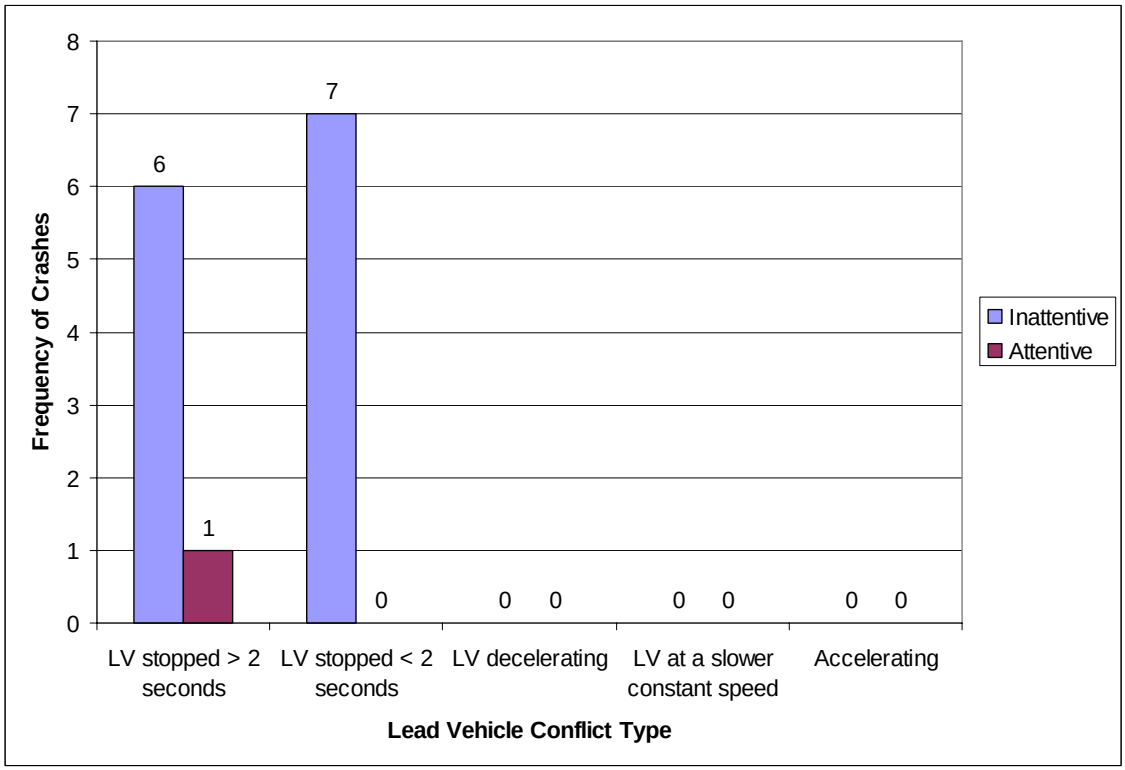


Figure RO.21. Frequency of crashes by driver attention and lead-vehicle kinematic scenario.

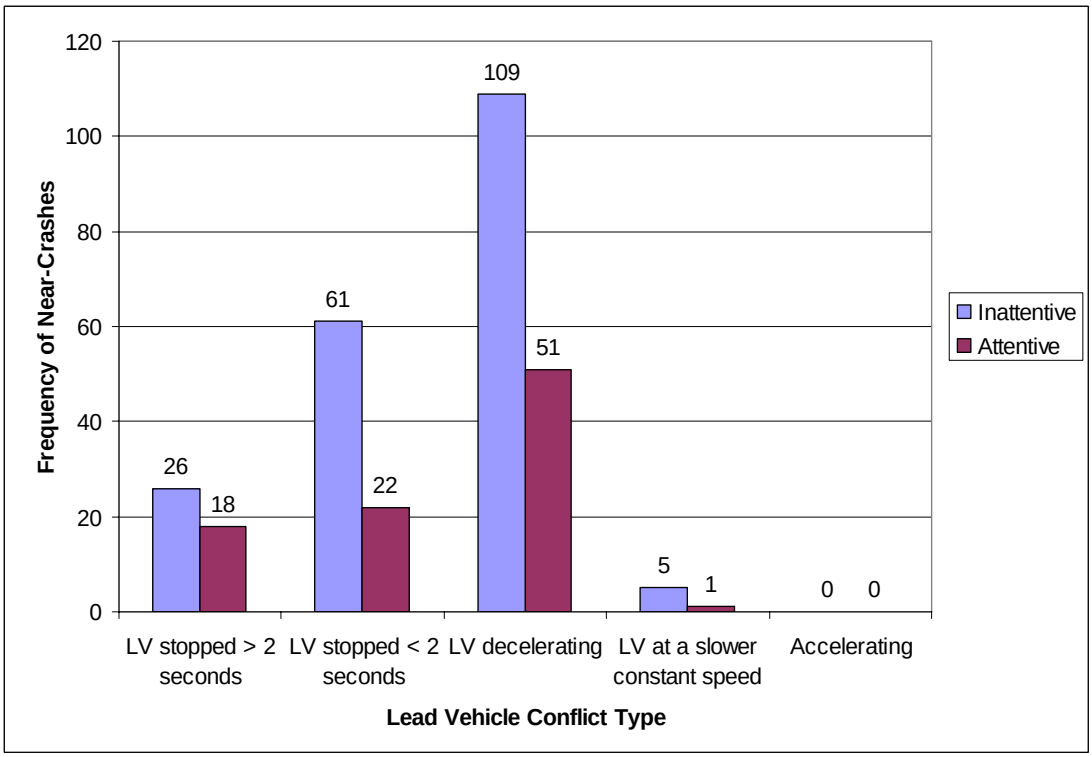


Figure RO.22. Frequency of near-crashes by driver attention level and LV scenario.

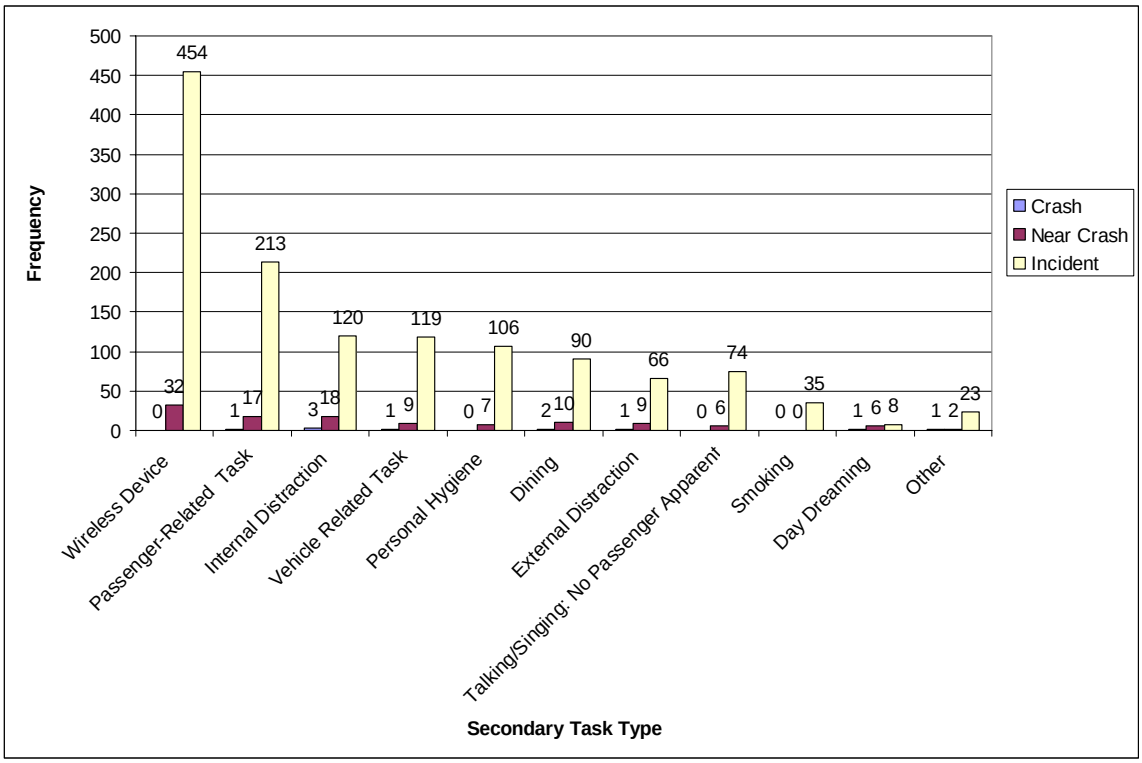


Figure RO.23. Frequency of secondary task inattention sources for lead-vehicle events.

GOAL 8: CHARACTERIZE THE REAR END SCENARIOS IN RELATION TO HEINRICH'S TRIANGLE

The purpose of *Goal 8* was to understand the relationship between the rates of crash, near-crash and incident events in order to potentially use near-crashes and incidents as *safety surrogates* in future empirical studies since they occur much more frequently than crash events. If proven reliable, such safety surrogates could be used in practice, for the first time, to predict the rate of crashes in a much more cost-effective manner than the collection of a statistically representative sample of crash events.

The premise behind Heinrich's triangle for both the original industrial safety application and the subsequent driving application is that the frequency of occurrence of "unsafe acts," or in the case of near-crashes and incidents, is related to the frequency of crashes. This has been shown to be the case in non-driving applications. For driving, the theory is that a crash most often is caused by a series of events including:

- a precipitating event;
- contributing factors; and
- the absence of a successful evasive maneuver.

Thus, in theory, joint probability models underlie the relative frequency of incidents, near-crashes, and crashes. That is, as we have operationally defined it, a precipitating event occurs, the associated presence of contributing factors determine whether it is responded to early (incident), late (near-crash, requiring evasive maneuver), or ineffectively (crash). The data in the

triangles, and the associated confidence limits, support the existence of such a possible underlying theory, although more data is needed to determine whether the associated frequencies are in fact stable to the point of having predictive value. The Heinrich's triangle for the lead-vehicle conflicts for this study is shown as Figure RO.24.

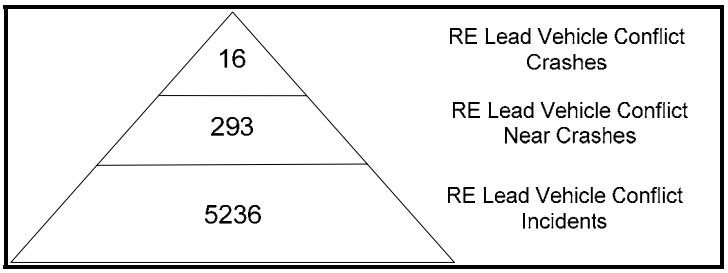


Figure RO.24. Heinrich's triangle for rear-end conflicts in the 100-Car Study.

Table RO.8 summarizes the rates of the differing severities of lead-vehicle conflicts and provides the 95 percent confidence limits modeled as a Poisson distribution for each category.

Table RO.8. Rates and confidence limits for each lead-vehicle conflict severity category.

	Count	Exposure per MVMT	Rate per MVMT	Variance (rate)	STD (rate)	Lower 95 percent CI for rate/MVMT	Upper 95 percent CI for rate/MVMT
RE Crashes	16	1.84	8.70	4.73	2.17	4.43	12.96
RE Near-Crashes	293	1.37	213.87	156.11	12.49	189.38	238.36
RE Incidents	5,236	1.37	3,821.90	2,789.71	52.82	3,718.38	3,925.42

These calculations suggest that lead-vehicle crashes occur at a rate of approximately 9 per MVMT within an approximate confidence interval of 4 to 12. This study observed 16 lead-vehicle crashes, which seems reasonable as the number of vehicle miles traveled is approaching 2 million VMT. Note the numerical stability, as indicated by the narrow confidence limits for the near-crash and incident data. Even though the number of crashes was fairly low, the crash confidence intervals are also approaching reasonable stability. While this data by itself is valuable, it indicates that such an approach may ultimately prove extremely useful when additional crash data is considered. That is, there is every indication that the approach of measuring less severe conflict surrogates may provide reasonable estimates of crash risk, particularly if a larger-scale naturalistic study can be conducted.

The types of analyses applied in this chapter were also conducted by Tijerina (2004) in his application of the Hazard Analysis Technique to data collected in the ADVANCE study (Dingus, 1997). Tijerina's application was unsuccessful; however, there are a couple reasons for this lack of success. First, as noted by Tijerina, the estimation of exposure was weak, since the ADVANCE database contained only 487 vehicle miles of data while the 100-Car Study collected 1.84 million vehicle miles.

Second, the crash data used in the analysis by Tijerina were taken from archival records for the preceding year based on 75.5 million vehicle miles. This is in contrast to the 487 vehicle miles for the near-crash and incident data. For the analyses in this chapter, the event data was taken

from the same database using the same drivers, meaning that the rates from the 100-Car Study are less prone to error. This inference seems to be confirmed through inspection of the confidence intervals for each rate estimate. As opposed to the Tijerina example, no rate estimates for which events were observed have confidence intervals that span zero.

GOAL 9: EVALUATE PERFORMANCE OF HARDWARE, SENSORS, AND DATA COLLECTION SYSTEM

Since a primary purpose of the 100-Car Study was to serve as a pilot for a larger scale (e.g., 5,000-car) study, one project goal was to understand the performance of the data collection system to assess the feasibility of such an undertaking. Thus, an analysis and utility of the various sensors and system components was performed.

Reliability was assessed in two ways. First major and catastrophic failures resulting in significant data loss were catalogued and analyzed. Second, minor failures, including the loss of a single data channel were analyzed. Catastrophic and major failure rates per sensor or subsystem are shown in Table RO.9. Sensors and subsystems not mentioned in the table did not exhibit any catastrophic or major failures. A total of 4,554 vehicle-weeks of data collection was used in the calculations. In addition, three weeks of downtime is assumed. This assumption is based on adding estimates for the time required to detect a failure (~1 week) and estimates for the time to perform a repair (~2 weeks). This estimate is somewhat conservative, since in many instances it took fewer than 3 weeks to detect and repair a fault, especially in the latter part of the study. Thus, the failure rates presented in this section represent a ceiling for the hardware used in the study.

Table RO.9. Catastrophic or major failure rates by sensor or subsystem.

Failing Sensor/Subsystem	Instances	Failure Rate (%)
Power Control Battery Backup	33	2.2
Acquisition Software	67	4.4
Remote Download	17	1.1
Real-time Video	22	1.4

Minor failure rates per sensor or subsystem are shown in Table RO.10. An assumption of three weeks downtime is used, along with a total data collection period of 4,554 vehicle-weeks. These 268 minor failures represent 804 vehicle-weeks of incomplete data. This means the overall minor failure rate (assuming independent failures and the downtime assumptions used before) was 17.7 percent. A total of 324,816.0 miles of data were incomplete, based on the assumed weekly mileage rate for the study of 404.0 miles per vehicle week. In some cases, this data could still be used in data reduction because a redundant source of data was available.

Table RO.10. Minor failure rates by sensor or subsystem.

Failing Sensor/Subsystem	Instances	Failure Rate (%)
Power Control Battery Back-up	6	0.4
Real-Time Video	97	6.4
Headway Detection	45	3.0
Vehicle Network	43	2.8
Lane Tracker	46	3.0
Remote Vehicle Tracking	8	0.5

Analyses were also conducted to determine the ability to develop *a priori* multivariate triggers to identify conflict events, including crashes, in a large-scale study. Recall that a 5,000-plus vehicle study will preclude the collection of continuous data due to the sheer volume of data and the necessity for video data to be present. Results for the analyses pointed to several conclusions that are relevant for a large-scale naturalistic data collection effort. First, crashes and near-crashes should be the focus of such an effort. Incidents are observed at a much higher rate than crashes and near-crashes; a total of 90.9 percent of all valid events were classified as incidents. Including incidents would likely overwhelm any data reduction effort for a large-scale study. Incidents are also closer in terms of kinematic signature to many invalid events than are crashes and near-crashes, making their discrimination more difficult.

Second, assuming that crashes and near-crashes are the focus of a large-scale study, tradeoffs concerning loss of valid events should focus on losing a minimal number of near-crashes. Based on the results of the discriminant analyses, changes in the sensitivity of the analysis had minimal effects on the number of crashes detected, but affected to a larger extent the number of near-crashes detected. Maximizing the number of near-crashes detected while minimizing the number of invalid events also tends to maximize the number of crashes detected.

Third, it seems that tailoring the triggering algorithms to particular individuals is a feasible partial solution to minimizing the number of invalid triggers collected, when it is combined with appropriately selected expected probabilities. This process was very effective in reducing the number of invalid events detected. Assuming that the 40th percentile longitudinal acceleration threshold is used to filter data, along with expected probabilities based on our current sample, the accuracy of a trigger algorithm could reach the levels shown in Table RO.11

Table RO.11. "Confusion matrix" showing hits, misses, false alarms, and correct rejections for the classification of near-crash events based upon multivariate trigger criteria.

		Event classified as:	
		Invalid	**Valid**
Event was:	**Invalid**	**79.8**	20.2
	Valid	28.1	**71.9**

It would also be expected that the majority of the valid events lost would be near-crashes, rather than crashes, given particular aspects of the crash event severity (e.g., longitudinal acceleration spikes) that make them easy to identify.

Achieving this tailoring process in a large-scale study would require some initial data collection on each participant's driving habits that would then be used to tailor the triggers for that driver,

which should always be the primary driver for the vehicle. This data collection period might be as short as a week, based on the data obtained for this study. While a small additional investment would be required to achieve this goal, the benefit gained by shortening the data reduction effort seems attractive.

The DAS used in this study purposefully contained a large number of sensors, some of which were redundant, with the goals of maximizing the level of redundancy within the system and obtaining a dataset that represented a nearly best-case scenario of data availability. This large number of sensors may not be needed for a larger-scale study. The events of interest may be more narrowly targeted or the magnitude of the data large enough that missing a few valid events is not as important as minimizing the number of invalid events that contaminate the dataset.

A larger-scale study would also magnify any system repair and/or maintenance needs. Thus, reducing the number of sensors and selecting sensors with low associated failure rates would be an important aspect of such an effort. Most of the sensors used in the data collection effort reported herein had very low failure rates, which will likely be even lower as technology progresses. The most problem-prone sensors were video and radar.

Given the advantages of video, however, it seems that its place as a sensor in a larger-scale study is necessary, although a smaller number of cameras might be acceptable. While the performance of the sensors in discriminating between valid and invalid events can be increased by data analysis methods, this increase is not large enough to warrant the elimination of the only method available for event verification.

The failure rate for radar was lower than for video. While there are problems with radar data, the radar units have to be carefully installed and they are usually damaged in crashes. The relative position and speed of leading traffic are important factors to consider for triggering to obtain valid events. Thus, despite the failure rate, the technology would be needed for a larger-scale study. Of course, if other technologies could sense the same data with a lower failure rate, they should be considered. At this time, however, no such technology exists at a reasonable price.

Other sensors, including accelerometers and gyros (for yaw rate), had negligible failure rates, undetectable for the current study. These sensors also provided data that proved very useful for valid event discrimination. These sensors should be included in the sensor suite for a large-scale study.

The triggers used in such an array of sensors would likely take values similar to those discussed in this goal, and the discrimination process using aggregate data would likely be equivalent. However, some of these triggers may become more stringent if higher accuracy sensors are used or if the data collection rate for some of the sensors is increased. The numbers suggested in this section for future use thus represent good starting values, although their performance should be tested within the final system in which they are included.

This page intentionally left blank.

CHAPTER 1: INTRODUCTION

BACKGROUND

There are two traditional approaches to collecting and analyzing human factors data related to driving. The first approach is to use data gathered through epidemiological studies (often collected on a national level). These databases, however, lack sufficient detail to be helpful for many applications, such as the development of countermeasure systems or the assessment of interactions between contributing factors that lead to crashes. The second approach, empirical methods, including newer, high-fidelity driving simulators and test tracks, is necessarily contrived and does not always capture the complexities of the driving environment or of natural behavior. For example, test subjects are often more alert and more careful in a simulation environment or when an experimenter is present in a research vehicle than when they are driving alone in their own cars. Thus, although empirical methods are very useful in other contexts, they provide a limited picture of the likelihood of a crash in a given situation or the potential reduction of that likelihood by a given countermeasure. State-of-the-art empirical approaches can only assess the relative safety of various countermeasures or scenarios. They cannot be used to predict the effect of a safety device or policy change on the crash rate.

Advances in sensor, data storage, and communications technology have led to the development of a hybrid approach to data collection and analysis that uses very highly capable vehicle-based data collection systems. This method of data collection has been used by some auto manufacturers since the introduction, several years ago, of electronic data recorders (EDRs). EDRs collect a variety of vehicular dynamic and state data that can be very useful in analyzing a crash. However, they currently lack sufficient measurement capability to assess many human factors issues.

Because of the shortcomings of traditional methods of data collection, it is becoming increasingly apparent that data collection in a "naturalistic" setting may be an effective approach for obtaining crash-related human factors data. Given the variability and complexity of driver behavior and performance, the random and rare nature of crashes, and the lack of adequate pre-crash data in today's crash record, it is especially important to collect real-world data that includes the crash experience and crash-relevant events in sufficient detail and depth. Such a dataset would make clear the conditions and driver behaviors that precipitate crashes as well as support the development and refinement of crash countermeasures.

In order to collect such a dataset, the National Highway Traffic Safety Administration (NHTSA) and the Virginia Department of Transportation (VDOT) contracted with the Virginia Tech Transportation Institute (VTTI) to conduct the "100-Car Naturalistic Driving Study." The study was a three-phased effort designed to meet the following objectives: Phase I, Conduct Test Planning Activities; Phase II, Conduct a Field Test; and Phase III, Prepare for Large-Scale Field Data Collection Effort. The Phase III effort will be completed and a report forthcoming prior to the end of the current contract. The large-scale field data collection effort is Phase IV, which is not being conducted under the current contract. This report describes the research methods, analyses, and results of Phase II.

Since Phase I included the foundation efforts for Phase II, a brief description of the tasks conducted under Phase I is first provided. A complete description of the 15 tasks, including task definition, methods, and results conducted under Phase I is provided in Neale et al. (2002).

PHASE I. CONDUCT TEST PLANNING ACTIVITIES

Task 1: Establish Intelligent Vehicle Initiative (IVI) Data Needs

The objective of Task 1 was to specify the details of the pre-crash and near-crash data to be gathered during the data collection phase. Pre-crash data involves all aspects of the driver's behavior and vehicle performance measures that occur prior to and leading up to a crash. Near-crash data involves the collection of driver behavior and vehicle performance data that occur in an event where the driver reacts appropriately so as to avoid a crash. Research questions involving analysis of pre-crash and near-crash data were generated and refined by the NHTSA Task Order Manager (TOM), interested NHTSA researchers, and other stakeholders in this project (Table 1.1).

Table 1.1. List of research goals organized by research category and subcategory.

General Research Category: Driver Behavior and Performance	
Subcategory: Crash/Near-Crash/Conflict Events	
Chapter 5, *Goal* 1	Classify and quantify contributing factors and dynamic scenarios involved in each event category.
Chapter 6, *Goal* 2	Operationally define a "near-crash" using quantitative measures.
Subcategory: Inattention Issues	
Chapter 7, *Goal* 3	Characterize driver inattention as it relates to incidents, near-crashes, and crashes.
Subcategory: Baseline Driving and Data Collection Issues	
Chapter 8, *Goal* 4	Characterize the differences in driving behavior and/or driver performance between: One week of field data and one year of naturalistic driving data, and One month in leased versus owned vehicle.
General Research Category: Distribution of Events	
Subcategory: Rear-End Event Analysis	
Chapter 9, *Goal* 5	Determine rear-end contributing factors and dynamic conditions. For each of the 4 rear-end (RE) lead-vehicle scenarios (stopped >2 s, decelerating, accelerating, moving at a slower constant speed), determine the frequency distribution of the following: crashes, near-crashes, and incidents: (i) per vehicle mile traveled (VMT); (ii) in relation to contributing factors; (iii) in relation to corrective actions; and (iv) in relation to transition events.
Chapter 10, *Goal* 6	Determine RE dynamics and precipitating factors -- specifically determine the frequency distribution for the following variables: Per VMT; initial kinematic condition; primary contributing factor. Crossed with these variables: Conflict with lead or following vehicle (crash, near-crash, incident); Conflict with lead or following vehicle when lead-vehicle changed lanes in front of subject vehicle; Conflict with lead or following vehicle when subject vehicle changed lanes behind lead vehicle; Conflict with lead or following vehicle when subject vehicle took corrective action.
Subcategory: Inattention Issues	
Chapter 11, *Goal* 7	Determine the distribution of inattention types for each RE lead-vehicle scenario (stopped >2 s, decelerating, accelerating, moving at a slower constant speed).
Subcategory: Baseline Driving and Data Collection Issues	
Chapter 12, *Goal* 8	Characterize each of the 4 RE lead-vehicle scenarios in relation to Heinrich's triangle.
General Research Category: Phase III Evaluations	
Subcategory: Vehicle Instrumentation	
Chapter 13, *Goal* 9	Evaluate the performance of the hardware, sensors, and data collection system used in data gathering (Phase II) in preparation for a future large-scale field study (Phase IV).
Subcategory: Data Reduction and Analysis	
Goal 10 (Separate Report)	Evaluate the performance of the data reduction plan, triggering methods, and data analysis in preparation for a future large-scale field study (Phase IV).

The general categories of research goals addressed driver behavior and performance, the distribution of collected driving events, and design of a future large-scale field study (which would be considered as Phase IV). The research goals lead to a set of candidate measures derived through a variety of methods including a literature review, a review of database variables

(e.g., police report form variables), and consultations with the TOM. The final set of 10 research goals are reported as separate chapters in this document.

Task 2: Develop Phase I Test Requirements

The Phase I test requirements were developed iteratively by VTTI with the cooperation of the TOM and stakeholders of this project. The primary Phase I test requirements addressed issues such as the number of cars to be instrumented, the number of camera views, the number of vehicle makes and models to be used, and the rate at which data was to be collected.

Task 3: Select Candidate Test Areas and Evaluate Crash Frequency Data

The objective of Task 3 was to determine the number of sites from which data could be collected, the rear-end crash frequency at various geographic locations, and the optimal location of the data collection site from the perspective of project resources. After consideration of these factors, the decision was made to collect data in the Washington, DC/Northern Virginia metropolitan area.

Task 4: Determine Crash Sampling Requirements

A goal of this study was to collect naturalistic data on approximately 10 rear-end crashes. In assessing the utility of the dataset, it was decided that continuous rather than triggered data would prove a greater value to the IVI program, other stakeholder organizations, and the development of the Phase IV protocol. From an operational and financial resource perspective, it was determined that 100 vehicles would be instrumented for continuous data collection. Although the number of crashes and other events to be captured could not be predicted with certainty, it was expected that this number of vehicles, driven by high-exposure drivers, would provide a sufficient number of crashes and other events (both general and of the rear-end type).

Task 5: Determine Driver/Vehicle Demographic Requirements

After a review of literature summarizing the driver factors that contribute to rear-end crashes, an ideal age and gender distribution was determined. Other recruiting factors, such as high-mileage drivers, roadway types traveled, and vehicle types were also determined. In addition, this task was conducted iteratively with Task 9: recruiting drivers, which provided further information for the driver selection criteria.

VTTI began determining the vehicle requirements by first establishing the primary criteria that should be considered in selecting vehicles. The task of choosing vehicle makes, model, and model years was conducted iteratively with Task 8: vehicle trade study, which provided further criteria for vehicle selection.

Task 6: Determine Near-Crash Statistical Power Requirements

In order to determine the near-crash statistical requirements, VTTI researchers reviewed four previous research studies that all used an instrumented vehicle in a natural driving environment. The frequency counts for crashes, near-crashes, and incidents were compared between these studies. The methods used to obtain these events were also compared, leading to the

development of estimates for the number of crashes, near-crashes, and incidents that could potentially be collected in this study.

Task 7: Conduct Trade Study – Research Design Parameters/ Sampling Rates/Formats Concept

A goal of 20 leased vehicles and 80 privately owned vehicles was set for the study. The final count was 22 leased vehicles and 78 privately owned vehicles. Research goals that NHTSA wanted to address by using these two groups of vehicles were: (1) the length of time required for the driver to adapt to an unfamiliar vehicle, and (2) the feasibility of using leased vehicles in the Phase IV large-scale data collection effort. Also, as one of the primary project goals was to collect pre-crash data for at least 10 rear-end crashes, the amount of data collected was paramount to the success of the study. These factors were incorporated into the experimental design.

Task 8: Conduct Trade Study to Determine Vehicle Types

Several factors were considered when determining the optimal vehicle types. The most critical factors included vehicle type, vehicle demographics, vehicle location, data collection system installation issues, information that could be obtained from the in-vehicle network, and the make and model requirements.

With regard to the model year of the privately owned vehicles, a review of model year revisions for each of the four selected vehicles indicated the following:

- The Toyota Camry was selected from the 1997 – 2001 model years. The Toyota Camry model design was static through 2001 with a new model in 2002.
- The Toyota Corolla was selected within model years 1993 – 2002.
- The Ford Explorer was chosen with model years 1995 – 2000 and model year 2001 (if manufactured by November 2, 2000).
- The Ford Taurus design was static over years 1996 – 1999, and a significantly different model was on sale for the years 2000 – 2002. In order to recruit sufficient drivers, both model sequences were used.

In addition, the Mercury Mountaineer is made on the same assembly line and has the same body style as the Ford Explorer; therefore, the Mercury Mountaineer could be included in the study for the same model years as the Ford Explorer. Likewise, the Mercury Sable is the same body style as the Ford Taurus, and could also be included in the list of potential vehicles.

Two additional makes and models were added as part of the leased vehicle portion of the fleet. Twenty vehicles were to be leased from the Virginia Tech Motor Pool - 10 model year 2002 Chevrolet Malibus and 10 model year 2002 Chevrolet Cavaliers. Obtaining the leased vehicles via the Motor Pool state contract was done to save project resources and reduce significant logistical problems (licensing, leasing agreements, etc.).

Task 9: Develop Participant Recruiting Specification

When developing the specifications for subject recruitment, the factors considered were: participant age, participant gender, vehicle types driven, the number of miles driven per year, and

location of either permanent residence or place of work. Participant age, gender, and annual mileage had important implications when considering the number of rear-end crashes expected to occur during the data collection period. The location of the participant's residence or work place was important when considering the difficulties of locating and downloading the data from the vehicles. Vehicle type was very important for private vehicle subjects as each vehicle had to be either be a Toyota (Camry or Corolla) or Ford (Taurus or Explorer). A hypothetical participant recruitment specification plan was developed based on the vehicle type, age group, and gender. Other issues addressed as part of this task were the example screening, classification questions, and issues of informed consent. Drivers who had been in many crashes were not given preferential treatment for inclusion in the study.

Task 10: Develop Test Data Collection Plan

Task 10 was a report requirement synthesizing Tasks 1 through 9. Comments were received from the contract sponsor and requested revisions were then made prior to continuation of this phase.

Task 11: Develop Test Reduction, Archiving, and Analysis Plan

The approach to data reduction for the Phase II study took advantage of an incident/near-crash data reduction method represented by Table 1.2, as well as current database information. Continuous data were collected and the incident/near-crash method was applied to the data (events were located in the dataset via optimized triggers that were determined through a sensitivity analysis). A data analysis plan was developed based on the research questions presented in Table 1.1.

The hardware aspects of data collection, back-up, and archiving were also described as part of this task, as well as the procedure for retrieving and organizing the data as they were obtained from the vehicles. A plan for long term data storage was also determined.

Table 1.2. Severity levels for the 100-Car Study.

Crash	Any contact with an object, either moving or fixed, at any speed, in which kinetic energy is measurably transferred or dissipated. Includes other vehicles, roadside barriers, objects on or off the roadway, pedestrians, cyclists, or animals.
Near-Crash	Any circumstance that requires a rapid, evasive maneuver by the subject vehicle (or any other vehicle, pedestrian, cyclist, or animal) to avoid a crash. A rapid, evasive maneuver is defined as steering, braking, accelerating, or any combination of control inputs that approaches the limits of the vehicle's capabilities. As a guide, subject vehicle braking greater than 0.5g or steering input that results in a lateral acceleration greater than 0.4g to avoid a crash, constitutes a rapid maneuver.
Crash-relevant conflict	Any circumstance that requires a crash avoidance response on the part of the subject vehicle or any other vehicle, pedestrian, cyclist, or animal that is less severe than a rapid evasive maneuver (as defined above), but greater in severity than a "normal maneuver" to avoid a crash. A crash avoidance response can include braking, steering, accelerating, or any combination of control inputs. A "normal maneuver" for the subject vehicle is defined as a control input that falls outside of the 99 percent confidence limit for control input as measured for the same subject.
Proximity Conflict	Any circumstance resulting in extraordinarily close proximity of the subject vehicle to any other vehicle, pedestrian, cyclist, animal, or fixed object when, due to apparent unawareness on the part of the driver, pedestrians, cyclists or animals, there is no avoidance maneuver or response. Extraordinarily close proximity is defined as a clear case in which the absence of an avoidance maneuver or response is inappropriate for the driving circumstances (e.g., speed, sight, distance, etc.).

Non-conflict event	Any event that increases the level of risk associated with driving, but does not result in a crash, near-crash, or conflict as defined above. Examples include driver control error without proximal hazards being present, driver judgment error such as unsafe tailgating or excessive speed, or cases in which drivers are visually distracted to an unsafe level.

Task 12: Development of Data Collection System Requirements

The results of Task 12 followed from the combined performance of Tasks 2 through 11. Additionally, Task 12 results were iterated and integrated with those of Tasks 13 and 15 to drive the Hardware/Software Design Specification. The data system requirements were categorized into four major areas:

1. Schedule Requirements;
2. General Design Requirements.;
3. Performance Requirements; and
4. Test Vehicle Profile.

Task 13: Review/Test of Technology/Sensor Alternatives

Tasks 13 and 15 were conducted in parallel to determine the most suitable hardware and software alternatives for each subsystem component. The data handling and software integration subsystems were addressed in Task 11. The remaining components were addressed in Tasks 13 and 15.

Task 14: Review/Test of Trigger Criteria Methods

Since it was decided early in the Phase I process that continuous data collection was desired, a triggered dataset was not needed. Instead, events in the dataset were to be located post hoc with editable triggers, which would result in a comprehensive database that could be filtered, scanned, sampled, and so forth, according to researchers' needs. The sensitivity analysis to determine the post hoc trigger levels was explained as part of Task 11.

Task 15: Trade Study Analysis of Hardware/Software Alternatives

Task 13 determined the available technologies and their relevant factors to meet the data collection system requirements determined as part of Task 12. Task 15 completed that effort. This task listed the subsystem component options considered in trade study analysis (as determined in Task 13), the evaluation performed to evaluate the component, the evaluation results, and the decision made for final component selection. The sensors and instruments to measure specified variables were discussed.

It may be noted that several variables were not meant to be collected through hardware. Driver classification and demographic variables were collected with questionnaires. Detailed vehicle information was collected prior to the study. Additional information on crashes was collected via police report forms.

This page intentionally left blank.

CHAPTER 2: METHOD FOR PHASE II – THE 100-CAR FIELD TEST

DRIVERS

One hundred drivers who commuted into or out of the Northern Virginia/Washington, DC metropolitan area were initially recruited as primary drivers to have their vehicles instrumented or receive a leased vehicle for this study. Drivers were recruited by placing flyers on vehicles as well as by placing newspaper announcements in the classified section. Drivers who had their private vehicles instrumented received $125 per month and a bonus at the end of the study for completing necessary paperwork. Drivers who received a leased vehicle received free use of the vehicle, including standard maintenance, and the same bonus at the end of the study for completing necessary paperwork. Drivers of leased vehicles were insured under the Commonwealth of Virginia policy.

As some drivers had to be replaced for various reasons (for example, a move from the study area or repeated crashes in leased vehicles), 109 primary drivers were included in the study. Since other family members and friends would occasionally drive the instrumented vehicles, data was collected on 148 additional drivers. Chapter 3 presents an exhaustive review of driver demographics.

THE 100-CAR DATA ACQUISITION SYSTEM

The 100-Car Study instrumentation package was designed and developed in-house by the VTTI Center for Technology Development. This system operated continuously after the system initialization period (or computer boot-up period, which required approximately 90 seconds after the ignition was turned on) until the driver turned the ignition off. Any commercial off-the-shelf components that were integrated into the instrumentation package are specifically noted in the following system description.

The core of the data acquisition system was a Pentium-based PC104 computer. The computer ran custom data acquisition software and communicated with a distributed data acquisition network. Each node on the network contained an independently programmable microcontroller capable of controlling or measuring a moderate number of signals. This system configuration maximized flexibility while minimizing the physical size of the system. Although capable of being expanded to 120 nodes, the vehicles were configured with 10 nodes. A schematic representation of the system appears in Figure 2.1.

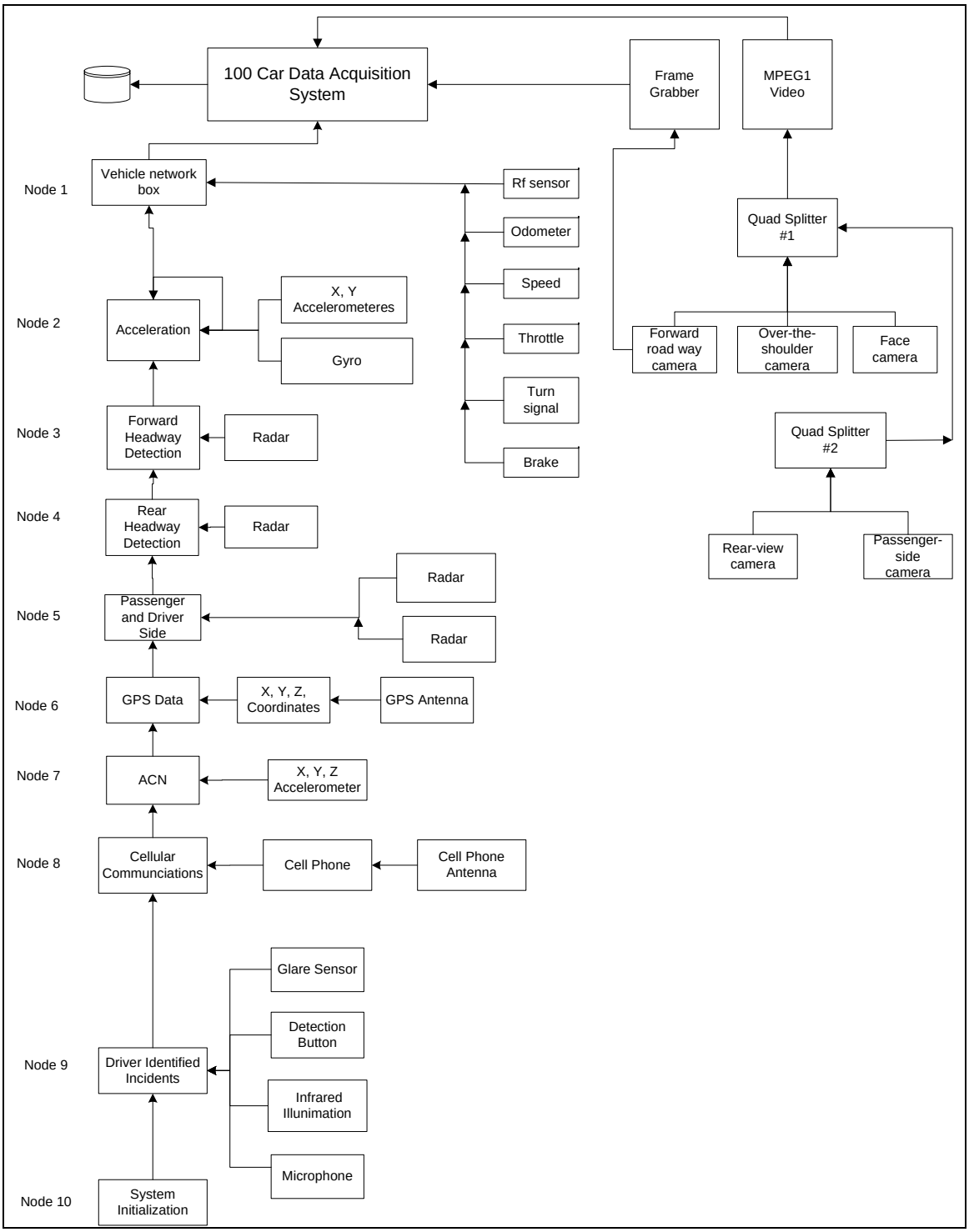


Figure 2.1. 100-Car Study data hardware collection system schematic.

This system of distributed data acquisition provided a very flexible and maintainable hardware data collection system. The main unit was mounted in the trunk under the "package shelf" (Figures 2.2 and 2.3). The vehicle network box was located under the front dashboard. The incident box was mounted above the rearview mirror. Wiring was run though the normal wire

chases on a vehicle to all the various network nodes, as well as to the cameras. All the microprocessor boards, including the firmware and data collection software, were developed at VTTI.

Figure 2.2. The main DAS unit mounted under the "package shelf" of the trunk.

Figure 2.3. The 100-Car Study DAS.

Node 1: Vehicle Network Box

This node was responsible for interfacing with the OBDII network in the vehicle. Various data elements were pulled off the network if they were available. Several sensors were hardwired such as the radio frequency sensor, the left turn signal, the right turn signal, and the brake light.

Node 2: Accelerometer Box

This node was responsible for collecting the lateral and longitudinal acceleration of the vehicle, along with the turning rate. MEMs based sensors were used.

Nodes 3-4: Headway Detection

These nodes were responsible for interfacing with an EATON VORAD EVT300 Doppler radar. Figure 2.4 shows the computer board for the node. The radars were mounted on the front and rear of the vehicles and were concealed behind plastic license plates (Figure 2.5).

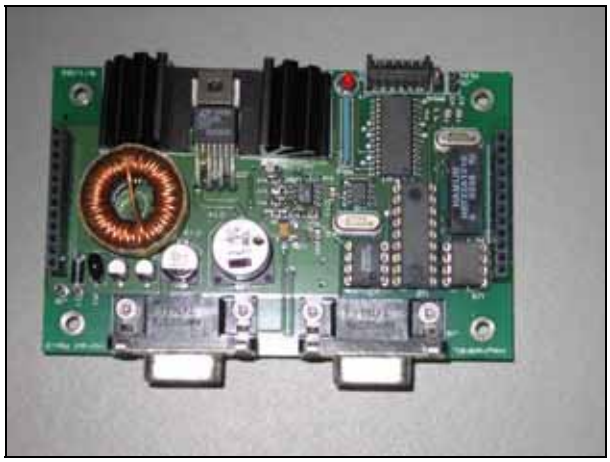


Figure 2.4. Computer board for the Vorad unit.

Figure 2.5. Radar unit mounted on the front of a vehicle, covered by a license plate.

Node 5: Side Obstacle Detection

These nodes were responsible for interfacing with a proprietary Doppler radar. These radars were capable of detecting targets at 30 ft and 180 degrees of span.

Node 6: GPS Data Node

This node was responsible for interfacing with a standard automotive GPS unit.

Node 7: Automatic Collision Notification

This node detected the possibility of a collision by sensing three accelerations. It would trigger a call to a dispatcher if it detected a crash.

Node 8: Cellular Communications

This node served as an interface between the computer and a standard cell phone. It was capable of receiving a call and connecting that call with the on-board computer, and likewise, the computer could call out.

Node 9: Incident Box

This node concentrated several data variables. It contained an incident pushbutton (shown mounted above the rear-view mirror in Figure 2.6) that the driver could press which would open an audio channel for the driver to verbally record an incident. It also housed the face camera, IR LEDs, and the glare sensor (shown mounted behind the rear-view mirror in Figure 2.7).

Figure 2.6. The incident pushbutton box mounted above the rearview mirror.

Figure 2.7. The mounting for the glare sensor behind the rearview mirror. Note the forward view camera as part of the same mounting assembly.

Node 10: System Initialization.

This node was responsible for qualifying the operating conditions, turning the computer on and off, and charging the cellular telephone backup battery. It also contained a watchdog functionality to maintain correct system operation, and a real-time clock for periodic system checkups.

Lane Tracking System

The lane tracking system incorporated a high resolution frame grabber and a full size image of the forward roadway. The data collection software ran an embedded version of a custom in-house machine-vision lane tracking system.

Video Data

There were 5 cameras located in the vehicle (Figure 2.8). One camera monitored the driver's face and the left side of the vehicle. A second camera monitored a 68° field of view (FOV) out the forward windshield. A third camera monitored a 68° FOV of the rear-view. The fourth camera monitored the passenger's side of the vehicle. Finally, the fifth camera monitored the driver's hands, instrument panel, and center console of the vehicle.

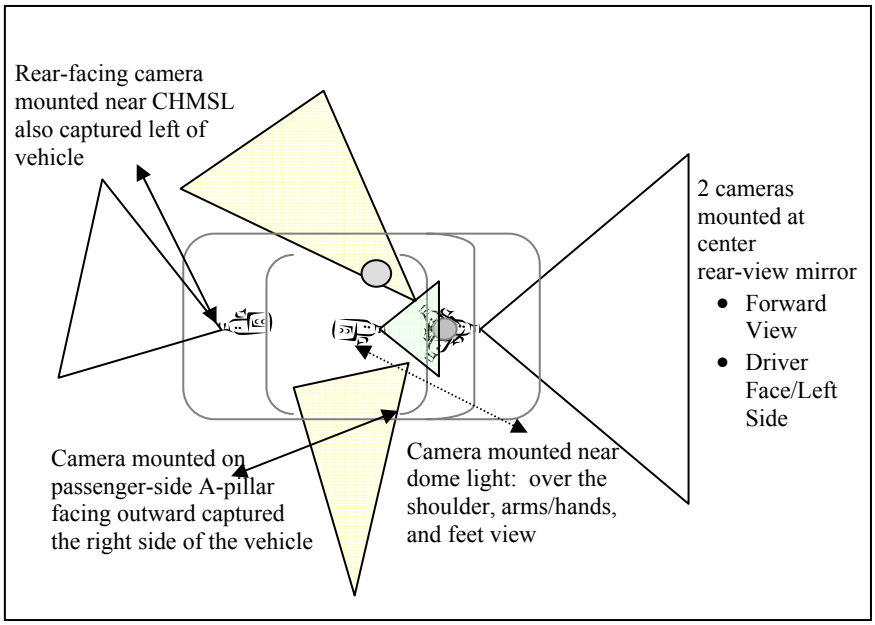


Figure 2.8. The 5 camera views recorded in the instrumented vehicle: (1) forward, (2) driver's face/left side of vehicle, (3) rear-facing, (4) over the driver's shoulder capturing the driver's hands and feet, the steering wheel, and the instrument panel, and (5) right side of vehicle.

Infrared lighting was used to illuminate the vehicle cab so that the driver's face and hands could be viewed on camera during nighttime driving. Figure 2.9 shows the placement and viewing angles of all 5 cameras in the quad-split image presented to allow data reductionists to monitor all 5 channels of video simultaneously.

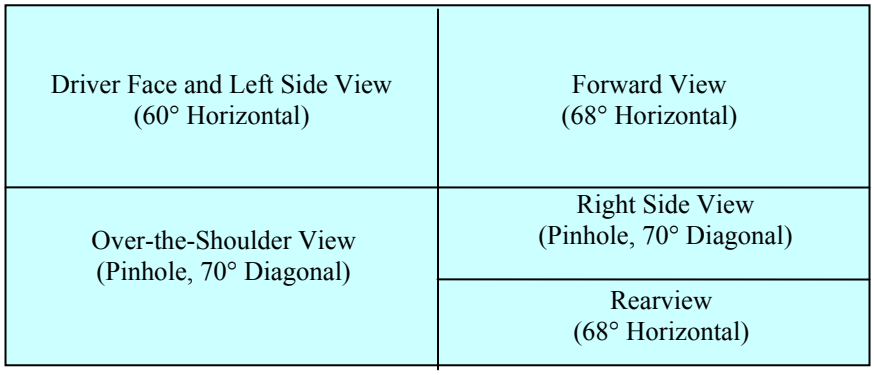


Figure 2.9. The double quad, split video image.

All video on board the 100-Car Study data collection system was compressed using MPEG 1 compression. This allowed greater storage of video on board the vehicle hard drives and required less server space to store the raw video data. While the initial data stream was recording at 30 Hz, the compression algorithm reduced the actual number of unique frames to approximately 7.5 frames per second (Figure 2.10).

Figure 2.10. A video image from the 100-Car Study data. The driver's face has been distorted to protect the driver's identity.

Driving Performance Data

Driving performance data were collected continuously and events were identified using specific values of driving performance dependent variables. Eleven main hardware sensor components were incorporated into the data collection system, as shown and described in Table 2.1 and depicted in Figure 2.1. In addition, relative lane position was derived using a combination of hardware on the instrumented vehicle and software written by VTTI computer programmers. This lane tracking system used machine vision based on input to the forward camera (prior to video compression). All data were stored in the data collection system in real-time.

Table 2.1. Description of Sensor Components.

Sensor Component	Description
Vehicle Network box	Collection of data directly from the in-vehicle network box. Some data includes vehicle speed, brake application, percent throttle, turn signal, etc.
Acceleration	Collection of lateral, longitudinal, and gyro.
Forward headway detection	Collection of radar data (range, range-rate, azimuth, etc.) to indicate the presence of up to 7 targets in front of the vehicle.
Rear headway detection	Collection of radar data (range, range-rate, azimuth, etc.) to indicate the presence of up to 7 targets behind the vehicle.
Side vehicle detection	Collection of radar data indicating the presence of a vehicle on the sides of the vehicle.
Global Positioning System	Collection of latitude, longitude, and horizontal velocity as well as other GPS-related variables.
Automatic Collision Notification System	High bandwidth collection of acceleration to detect a severe crash.
Cellular communications	Communication system designed for vehicle tracking and system diagnostics.
Driver Identified Events/Glare sensor	Collection of lux value (for night-time conditions only) as well as event button.
System Initialization	Overall system operation.

DATA COLLECTION, ARCHIVAL AND STORAGE

Demographic and Questionnaire Data

Prior to the installation of the data collection system in the participant's vehicle or acquisition of a leased vehicle, each participant met with a VTTI researcher at the UVA/VT Northern Virginia Center in Falls Church, VA. During this meeting, a VTTI researcher:

- Obtained informed consent from the private-vehicle or leased-vehicle participant, and explained that a Certificate of Confidentiality had been obtained from the National Institute of Mental Health for the participant's protection.
- Explained that the study was investigating traffic in northern Virginia.
- Explained the logistics of data collection system installation and maintenance.
- Asked the participant to agree to a vision and hearing exam.
- Asked the participant to complete questionnaires and take two computer-based tests.

The tests and questionnaires, as well as whether these were completed prior to or after data collection, are listed in Table 2.2. Full text versions of the informed consent form, tests, and questionnaires are located in Appendix A.

Table 2.2. Description of all tests and questionnaires administered to study participants.

Test/Questionnaire	Test Type	When Administered	Brief Description
1. Visual Acuity Test	Performance test using verbal report	Before data collection	Used the Snellen Eye Chart to test driver's visual acuity.
2. Audiogram Air Conduction Test	Examination using an audiometer	Before data collection	Assessed hearing levels at a frequency range of 125-8000 Hz.
3. Medical Health Assessment	Questionnaire	Before data collection	Obtained any information on prior health problems that may relate to driving performance.
4. Walter Reed Army Institute of Research Preliminary Sleep Questionnaire	Questionnaire	Before data collection	Measured and recorded subject's sleep habits and problems that may cause drowsiness.
5. Dula Dangerous Driving Index	Questionnaire	Before data collection	Classified driver's level of aggressive driving behavior.
6. Driver Stress Inventory	Questionnaire	Before data collection	Used a 10-point Likert Scale to obtain information about driver's general attitudes toward driving on a variety of roadways and in traffic congestion.
7. Life Stress Inventory	Questionnaire	Before and after data collection	Obtained information about the types of stress and changes that the subject may have experienced in the past year to determine the risk level for illness.
8. NEO FFI (Neuroticism Extraversion Openness Five Factor Model)	Questionnaire	Before data collection	Measured the five dimensions of normal personality: neuroticism; extraversion; openness; agreeableness; and conscientiousness.
9. Way Point	PC-based performance test	Before data collection	Used to identify drivers who may be at high risk for crashes by measuring their information processing speed and aptitude for vigilance.
10. Useful Field of View (UFOV)	PC-based performance test	Before data collection	Used to measure a driver's risk for crash involvement by using the driver's central vision and processing speed, divided attention, and selective attention.
11. Debriefing Questionnaire	Questionnaire	After data collection	List of questions collecting information on driver's recollections about events that occurred during the last year, seat belt use, alcohol use, etc.
12. Driver Demographic Information	Questionnaire	Before data collection	List of questions collecting information on driver's age, gender, level of education, occupation, etc.
13. Driving History	Questionnaire	Before data collection	List of questions collecting information on driver's traffic violations and accident history, type, etc.
14. Post-Crash Interview Form	Interview questionnaire	In the event of a crash	Used to collect driver's description of crash
15. Seatbelt	Questionnaire	Before data collection	Assessed seatbelt use and attitudes toward seatbelt use.

Instrumentation

The instrumentation of vehicles was orchestrated by the VTTI's Center for Technology Development. The 22 leased vehicles were instrumented at VTTI and the 78 private vehicles were instrumented by a Northern Virginia company. Using a phased approach, it took four months to get all 100 vehicles on the road. Since the data collection system automatically powered on and off, data collection began on each vehicle as soon as it was instrumented.

Data Retrieval and Storage

To collect the data from the experimental vehicles, "chase vehicles" were used to track the vehicle, go to the location, and download data. The chase vehicle drivers "called" the vehicle using a cellular telephone and laptop configuration. In-house software then displayed a map showing icons for the chase vehicle and experimental vehicle locations. The chase vehicle driver then drove to the location of the instrumented vehicle and downloaded the data from the experimental vehicle (downloading required a data transfer cable connected to an outlet near the rear license plate of the instrumented vehicle, which was connected to a data storage device). After each download, data integrity was verified. Data were again duplicated in Northern Virginia onto DVDs, one copy was sent to VTTI and the other copy was kept in Northern Virginia.

As the data arrived at VTTI, the triggering software was run on each DVD (see "Data Reduction') and the resulting relevant event epochs were saved. Event epochs were copied and saved on the networked attached storage server (NAS) at VTTI. The remainder of the video and raw data contained on the DVD remained on the DVD.

Once the triggered data were copied to the NAS at VTTI, the data were deleted from the experimental vehicle hard drive using in-house software. Once the data arrived at VTTI a fourth copy was created on the NAS before the on-board data were deleted. The purpose of this detailed duplication and storage scheme was to maintain a minimum of two data copies at all times.

PROCEDURE FOR DATA REDUCTION

Sensitivity Analysis

As stated previously, data were collected continuously to optimize the trigger criteria values after driving performance data were collected. If the triggers had been set prior to data collection, valuable events may have been lost without any method of recovery. One method of efficiently establishing trigger criteria is to perform a sensitivity analysis.

Figure 2.11 shows the data reduction plan in a flow chart format. Raw data from the vehicles was saved on the NAS at VTTI until approximately 10 percent of the data expected to be collected for the entire study was stored on the NAS. At that time, a sensitivity analysis was performed to establish post-hoc trigger criteria.

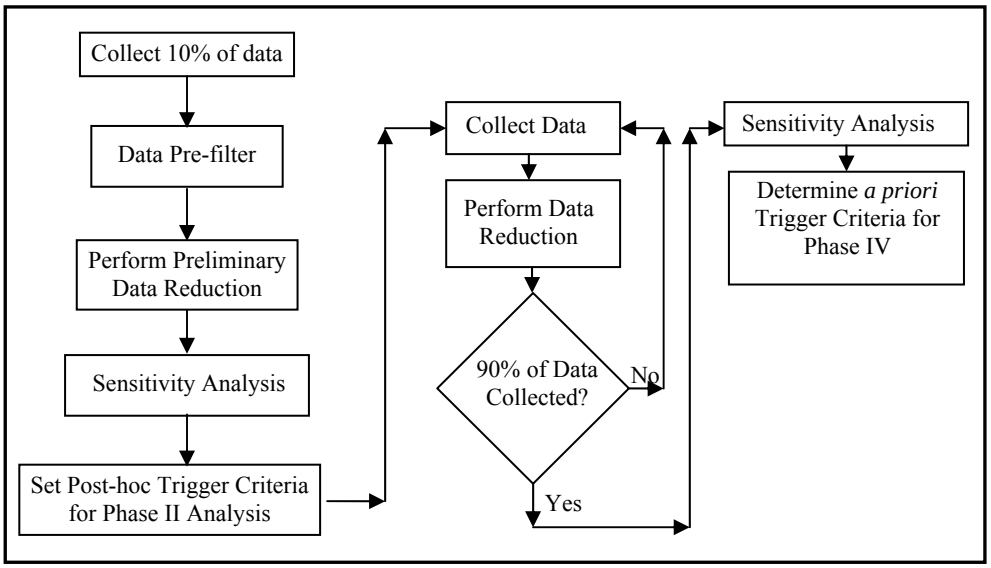


Figure 2.11. Flow chart of the data reduction process.

The sensitivity analysis was conducted by making iterative adjustments to the trigger values to ensure that most of the valid events were identified with only a few invalid events also being identified. The list of dependent variables ultimately used as event triggers is presented in Table 2.3.

Table 2.3. Dependent variables used as event triggers.

Trigger Type	Description
1. Lateral Acceleration	• Lateral motion equal to or greater than 0.7 g.
2. Longitudinal Acceleration	• Acceleration or deceleration equal to or greater than 0.6g. • Acceleration or deceleration equal to or greater than 0.5 coupled with a forward TTC of 4 seconds or less. • All longitudinal decelerations between 0.4g and 0.5g coupled with a forward TTC value of ≤ 4 seconds and that the corresponding forward range value at the minimum TTC is not greater than 100 ft.
3. Event Button	• Activated by the driver by pressing a button located on the dashboard when an event occurred that he/she deemed critical.
4. Forward Time-to-Collision	• Acceleration or deceleration equal to or greater than 0.5 coupled with a forward TTC of 4 seconds or less. • All longitudinal decelerations between 0.4g and 0.5g coupled with a forward TTC value of ≤ 4 seconds and that the corresponding forward range value at the minimum TTC is not greater than 100 ft.
5. Rear Time-to-Collision	• Any rear TTC trigger value of 2 seconds or less that also has a corresponding rear range distance of ≤ 50 feet AND any rear TTC trigger value in which the absolute acceleration of the following vehicle is greater than 0.3g
6. Yaw rate	• Any value greater than or equal to a plus AND minus 4 degree change in heading (i.e., vehicle must return to the same general direction of travel) within a 3 second window of time.

A sensitivity analysis was performed by setting the trigger criteria to a very liberal level, reducing the chance of a missed valid event to a minimal level while allowing a high number of

invalid events (false alarms) to be identified (see Figure 2.12). Data reductionists then viewed all of the events produced from the liberal trigger criteria and classified each event as valid or invalid. The number of valid events and invalid events that resulted from this baseline setting was recorded.

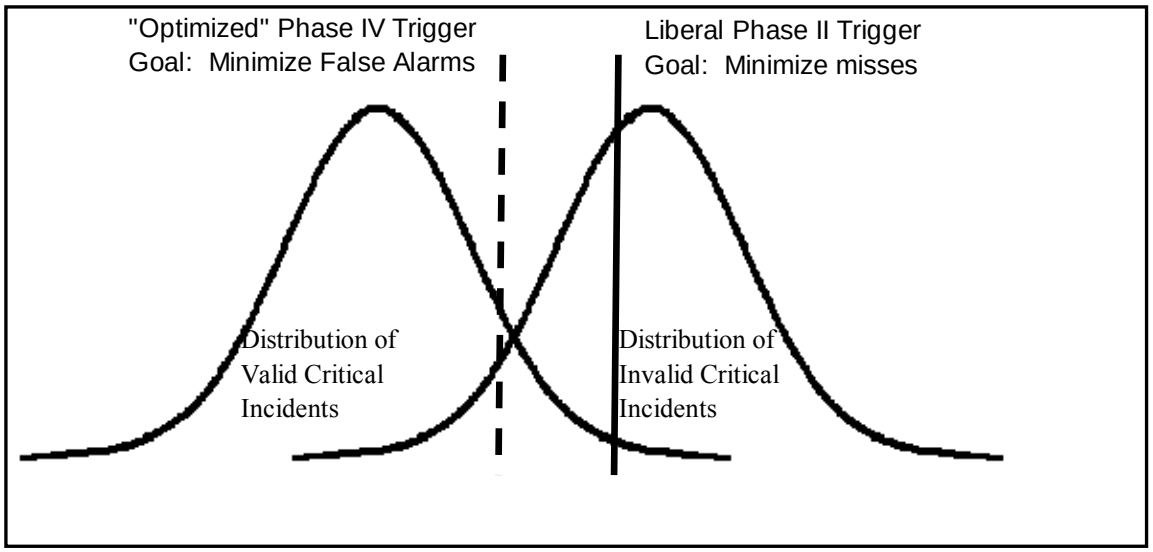


Figure 2.12. Graphical depiction of trigger criteria settings for Phase II and Phase IV using the distribution of valid events. Note that this distribution and criterion placement is unique for each trigger type.

The trigger criteria for each dependent variable was then set to a slightly more conservative level and the resulting number of valid and invalid events was counted and compared to the first frequency count. The trigger criteria were made more and more conservative and the number of valid and invalid triggers counted and compared until an optimum trigger criteria value was determined (a level which results in a minimal amount of valid events lost and a reasonable amount of invalid events identified). The goal in this sensitivity analysis was to obtain a miss rate of less than 10 percent and a false alarm rate of less than 30 percent.

Based on data from past VTTI studies, it was originally hypothesized that as many as 26 crashes, 520 near-crashes, and over 25,000 incidents (crash-relevant conflicts and proximity conflicts) would be collected; however many of these early estimates were based on long-haul truck driving data. It was soon discovered, after the sensitivity analysis process began, that the variability in light vehicle drivers" braking, acceleration, and steering behavior is much larger than with truck drivers. It is likely that this is due to differences in vehicle dynamics and the more uniform driving skill of commercial truck drivers.

Given the large variability in light vehicle driving performance, the sensitivity analysis proved to be challenging. VTTI researchers determined that the best option was to accept a very low miss rate while accepting a fairly high false alarm rate to ensure that few valid events were missed. This resulted in viewing over 110,000 events in order to validate 10,548 events. The distribution of the total number of reduced events by severity is shown in Table 2.4.

Table 2.4. The total number of events reduced for each severity level.

Event Severity	Total Number
Crash	69 (plus 13 without complete data)
Near-Crash	761
Incidents (Crash-relevant Conflicts and Proximity Conflicts)	8,295
Non-Conflict Events	1,423

Once the trigger criteria were set for Phase II, data reductionists watched 90-second epochs for each event (one minute prior to and 30 seconds after), reduced and recorded information concerning the nature of the event, driving behavior prior to the event, the state of the driver, the surrounding environment, etc. The specific variables recorded in the data reduction process are described in detail in the data reduction software framework section of this chapter.

Recruiting and Training Data Reductionists

Based upon past experience, it was estimated that reductionists would be able to reduce an average of 4 events per hour. Eleven data reductionists were recruited by posting flyers and notices to various graduate student listserves on the Virginia Tech campus. The data reduction manager interviewed, hired and trained the data reductionists on how to access the data from the server and operate the data reduction software, and provided training on all relevant operational and administrative procedures (approximately 4 hours of training). The manager gave each data reductionist a data reduction manual to guide them in learning the software and reduction procedures. All analyst trainees practiced data reduction procedures with another trained analyst prior to reducing data independently. After each trainee felt comfortable with the process, the trainee worked alone under the supervision of the data reduction manager. Once the trainee and manager felt confident of the analyst's abilities, the analyst began working independently, with "spot check" monitoring from the project leader and other reductionists. The data reductionists were responsible for analyzing a minimum number of events per week, and were required to attend weekly data reduction meetings to discuss issues that arose in data reduction.

The data reductionists performed two general tasks for this project. On the first 10 to 15 percent of the data, they performed a preliminary data reduction task in which they viewed events to determine whether the event was valid or invalid and to determine the severity of the event. After the trigger criteria for Phase II was set using the results from the sensitivity analysis, the data reductionists then validated the data, determined severity, and performed a full data reduction. For the full data reduction, they recorded all of the required variables (discussed below) for the event type. To ascertain severity of the event, reductionists used the decision tree, as shown in Figure 2.13.

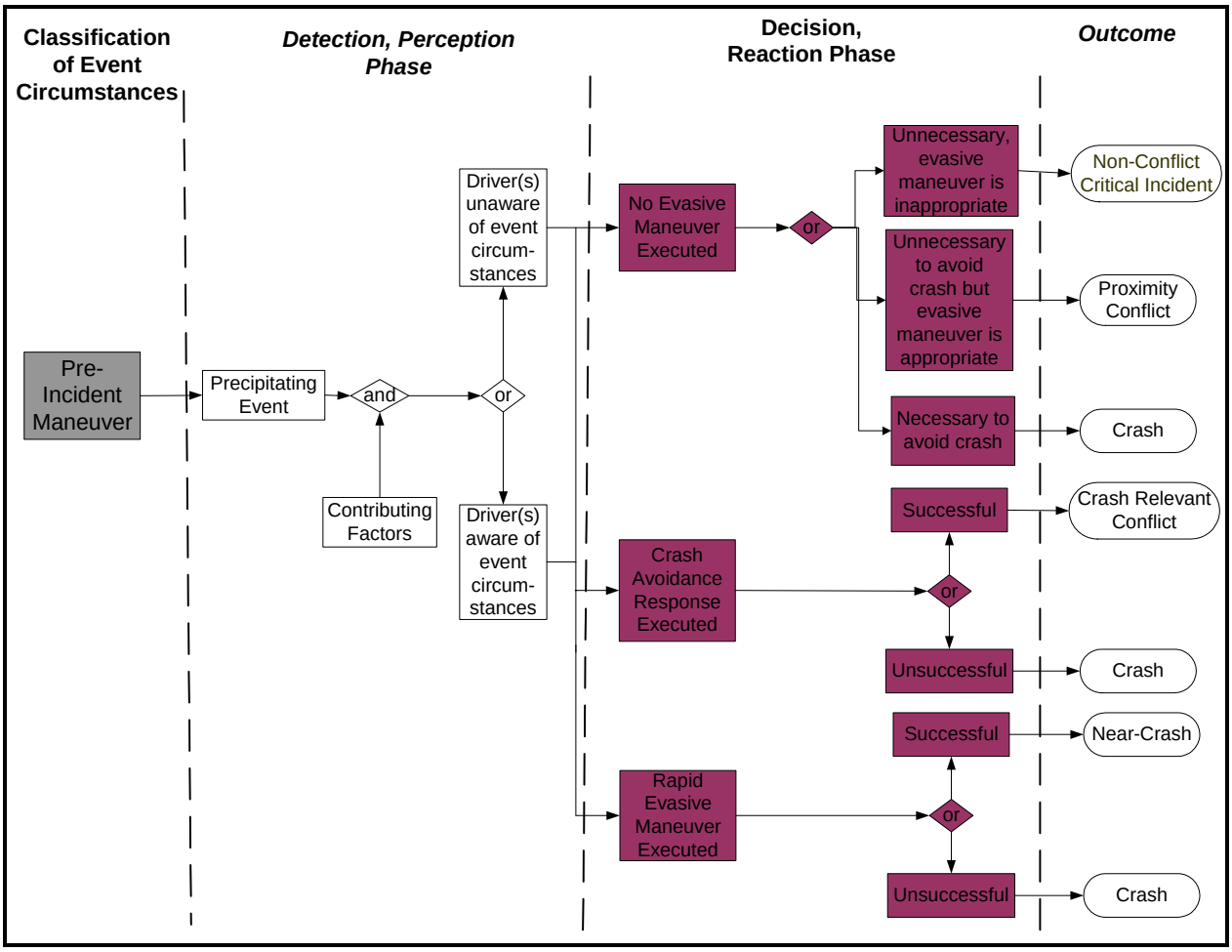


Figure 2.13. Decision tree used to classify event severity.

Data Reduction Software Framework

The data reduction framework was developed to identify various driving behavior and environmental characteristics for four levels of event severity: crashes; near-crashes; crash-relevant conflicts; and proximity conflicts. The variables recorded were selected based upon past instrumented vehicle studies (Hanowski et al., 2000; Dingus et al., 2002), national crash databases (General Estimates System and Fatality Accident Reporting System), and questions on Virginia State Police Accident Reports. Using this technique, the reduced database can be used to directly compare crash data from GES and FARS to those crashes, near-crashes, and incidents (crash-relevant conflicts and proximity conflicts) identified in this dataset.

The general method for data reduction was to have trained data reductionists view the video data and record the battery of variables for all valid events. The data reduction manager and project manager performed all data reduction on the near-crashes and crashes. Varying levels of detail were recorded for each type of event. Crash-relevant conflicts and proximity conflicts have the least amount of information recorded and near-crashes and crashes have the most information recorded. A total of four areas of data reduction were recorded for each event type. These four

areas include: vehicle variables; event variables; environmental variables; and driver state variables. Table 2.5 defines each area of data reduction, provides examples, and describes additional features of the data reduction. The complete list of all variables reduced during data reduction is shown in Appendix B.

Table 2.5. Areas of data reduction, definition of the area, and examples.

Area of Data Reduction	Definition	Example
Vehicle Variables	All of the descriptive variables including the vehicle identification number, vehicle type, ownership, and those variables collected specifically for that vehicle (VMT).	Vehicle ID, Vehicle type, Driver type (leased or private), and VMT.
Event Variables	Description of the sequence of actions involved in each event, list of contributing factors, and safety or legality of these actions.	Nature of Event/ Crash type, Pre-event maneuver, Precipitating Factors, Corrective action/Evasive maneuver, Contributing Factors, Types of Inattention, Driver impairment, etc.
Environmental Variables	General description of the immediate environment, roadway, and any other vehicle at the moment of the incident, near-crash, or crash. Any of these variables may or may not have contributed to the event, near-crash or crash.	Weather, ambient lighting, road type, traffic density, relation to junction, surface condition, traffic flow, etc.
Driver's State	Description of the instrumented vehicle(s) driver's physical state.	Hands on wheel, seat belt usage, fault assignment, eyeglance, PERCLOS, etc.
Driver/Vehicle 2	Description of the vehicle(s) in the general vicinity of the instrumented vehicle and the vehicle's action.	Vehicle 2 body style, maneuver, corrective action attempted, etc.
Narrative	Written description of the entire event.	
Dynamic reconstruction	Creation of an animated depiction of the event.	

Data Reduction Inter- and Intra-Rater Reliability

Training procedures were implemented to improve both inter- and intra-rater reliability, given that data reductionists were asked to perform subjective judgments on the video and driving data. Reliability testing was then conducted to measure the resulting inter- and intra-rater reliability.

First, data reductionist managers performed spot checks of the reductionists' work, monitoring both event validity judgments as well as recording all database variables. Reductionists also performed 30 minute's worth of spot-checks of their own or other reductionists' work every week. This was done to ensure accuracy but also to allow reductionists the opportunity to view other reductionists' work. It was anticipated that this would encourage each reductionist to modify their own work and to improve consistency in decision-making techniques across all reductionists. Mandatory weekly meetings were held to discuss issues concerning data reduction techniques. Issues were usually identified by the spot-checking activities of the reductionist managers and the reductionists, or specific difficult events that the reductionists had encountered. These meetings provided iterative and on-going reduction training throughout the entire data reduction process.

To determine how successful these techniques were, an inter- and intra-rater reliability test was conducted during the last three months of data reduction. Three reliability tests were developed (each containing 20 events) for which the reductionist was required to make validity judgments. Three of these 20 events were also fully reduced. Three of the test events on Test 1 were repeated on Test 2 and 3 other events were duplicated between Tests 2 and 3 to obtain a measure of intra-rater reliability.

Using the expert reductionists' evaluations of each epoch as a "gold" standard, the proportion of agreement between the expert and each rater was calculated for each test. The measures for each rater for each testing period, along with a composite measure, can be found in Table 2.6.

Table 2.6. Percentage agreement with expert reductionists.

Rater	Test 1 Percent	Test 2 Percent	Test 3 Percent
1	78.3	87.5	91.3
2	65.2	70.8	78.3
3	100	91.7	95.7
4	100	91.7	87.0
5	100	83.3	87.0
6	95.7	87.5	91.3
7	91.3	87.5	91.3
8	91.3	91.7	91.3
9	95.7	70.8	91.3
10	95.7	91.7	87.0
11	95.7	87.5	100
12	78.3	87.5	87.0
13	87.0	83.3	96.0
14	78.3	83.3	91.3
	Average (across all tests)	88.4	

The Kappa statistic was also used to calculate inter-rater reliability. Although there is controversy surrounding the usefulness of the Kappa statistic, it is viewed by many researchers as the standard for rater assessment (e.g., Cicchetti and Feinstein, 1990). The Kappa coefficient ($K = 0.65$, $p < 0.0001$) indicated that the association among raters is significant. While the coefficient value is somewhat low, given the highly subjective nature of the task, the number of raters involved, and the conservative nature of this statistic, the Kappa calculation probably errs on the low side.

A tetrachoric correlation coefficient is a statistical calculation of inter-rater reliability based on the assumption that the latent trait underlying the rating scale is continuous and normally distributed. Based on this assumption, the tetrachoric correlation coefficient can be interpreted in the same manner as a correlation coefficient calculated on a continuous scale. The average of the pair-wise correlation coefficients for the inter-rater analysis is 0.86. The coefficients for the intra-rater analysis were extremely high with 9 raters achieving a correlation of 1.0 among the three reliability tests and 5 raters achieving a correlation of 0.99.

Given these three methods of calculating inter-rater reliability, it appears that the data reduction training coupled with spot-checking and weekly meetings proved to be an effective method for achieving high inter- and intra-rater reliability.

Database Creation

All of the data analyses in this report are based on (1) driving performance data derived from the raw data collected on-board the vehicles, (2) reduced data resulting from the event analysis, and (3) subjective questionnaires filled out by subjects pre- and post-data collection. These data were copied, created, or edited into MySQL databases and linked using identification codes (i.e., vehicle or epoch identification numbers). Using these databases, it was then possible to identify, for example, the number of near-crashes and crashes for male drivers under age 24 compared to males drivers over age 45 for relationship to crash involvement.

BY THE NUMBERS – TOP LEVEL PROJECT STATISTICS

The final top-level statistics for the 100-Car Study are provided in Table 2.7. Note that 109 primary drivers drove 100 vehicles, of which 78 were personal vehicles, and 22 were leased vehicles. More than 100 primary drivers were used because some drivers dropped out of the study and others were replaced for various reasons. Altogether there were 241 total drivers (primary drivers plus secondary drivers). Over 6 terabytes of data were collected and stored on over 1,300 DVDs. Altogether, there were 82 crashes. Of those, complete data were available for 69. Also, of the 82 crashes, 49 were low g events, such as struck or ran over curb, median, parking blocks, or small animal). There were 761 near-crashes and over 8,000 incidents.

Table 2.7. Top-level 100-Car Study statistics.

Parameter	Statistic
Participants:	109 primary drivers 241 total drivers
Vehicles:	78 personal, 22 leased
Miles driven:	2,025,000
Hours of driving data collected:	47,382.65
Average speed:	29 mph
Overall duration of data collection in months:	18.5
Amount of data in terabytes:	6.4 TB
Amount of data in DVDs:	1,361 DVDs
Crashes (see Table 2.8):	82 (69 with complete data)
Near-Crashes:	761
Incidents:	8,295

The 82 crashes are summarized in Table 2.8 in terms of crash type and whether or not the crash was reported to police. The most common crash types were Rear-End Striking (29 percent of total) and Rear-End Struck (25 percent of total). Single Vehicle Run-Off-Road was the third most common at 18 percent of the total. The other main contributor to the overall total was Backing, at 13 percent of the total; however, none of these Backing crashes were police reported, while at least some of all the other most common types were police reported.

Table 2.8. Summary of crashes.

Type	Police Reported*	Number	Percentage of Total
Single Vehicle Run-off-Road	No	29	35%
Single Vehicle Run-off-Road	Yes	3	4%
Rear-end, striking	No	11	13%
Rear-end, striking	Yes	5	6%
Rear-end, struck	No	12	15%
Rear-end, struck	Yes	3	4%
Backing	No	8	10%
Backing	Yes	0	0%
Left Turn Across Path	No	0	0%
Left Turn Across Path	Yes	2	3%
Sideswipe	No	2	3%
Sideswipe	Yes	0	0%
Lane Change	No	1	1%
Lane Change	Yes	0	0%
Hit by object	No	1	7%
Hit by object	Yes	0	0%

* Crashes were counted as non-police-reported when this was not known.

The ratios of police reported crashes to non-police reported crashes varied considerably depending on crash type. For example, none of the Backing crashes were police reported. The overall ratio of non-reported to reported crashes was 2.9 to 1 (i.e., there were 2.9 non-reported crashes for every reported crash). Several categories of crashes were all police reported (Rear-End Striking and Struck, Left Turn across Path, and Lane Change), while other categories were not reported at all (Backing, Sideswipe, and Hit By Object). Categories in which some crashes were reported and some crashes were not reported included Single Vehicle Run-Off-Road, Rear-End Striking, and Rear-End Struck. There were 38 crashes in the three most common crash type categories, and the ratio of non-reported to reported crashes for these three categories was 3.2 to 1.

Table 2.9. Ratios of non-police-reported to police reported crashes.

Category	Numbers
Overall ratio of non-reported crashes to reported crashes	2.7 : 1
Non-police reported or unknown if police reported	41
Known police reported	14
Ratio by crash type:	
Single Vehicle Run-off-Road	2 : 1
Rear-end, striking	2.2 : 1
Rear-end, struck	12 : 1
Rear-end, struck & striking	All police reported
Backing	All non-police reported
Left Turn Across Path	All police reported
Sideswipe	All non-police reported
Lane Change	All police reported
Hit by object	All non-police reported
Three most common crash types (37 crashes)	3.2 : 1

When comparing these statistics to the expected crash rates as cited in the 100-Car Phase I Report, there were significantly more rear-end crashes than were expected. Using Crash Rate Calculation 4 from the Phase I Report, the calculation was based on the following sources or assumptions:

1. Northern Virginia/ Washington, DC, metropolitan area crash rate statistics.
2. Assuming 2.88 MVMT.
3. Biasing the sample towards younger drivers.
4. Higher crash rate in an urban area.

These sources and assumptions suggested that there would be data for 6.94 police-reported rear-end crashes and potentially 10 rear-end crashes accounting for all of the non police-reported crashes.

The final numbers for rear-end crashes were as follows:

- 31 total striking and struck rear-end crashes (reported or identified in the 100-Car Study database).
- 8 police-reported rear-end striking and struck rear-end crashes.
- 23 crashes were non-police reported.

Note that 1.4 MVMT were collected during the 100-Car Study and that the driver sample was only slightly biased toward younger drivers with 50 percent of the drivers under age 35. These results suggest that police-reported crash statistics greatly underestimate the actual number of crashes that occur.

The numbers of primary drivers involved in incidents of various types is shown in Table 2.10. The weekly dataset developed for Chapter 8, *Goal* 4 was used in this analysis. There were data available for 107 0f the 109 primary drivers in this dataset. It can be seen that over 35 percent of drivers were involved in at least one crash, while over 80 percent experienced at least one near-crash and over 90 percent were involved in at least one incident. Table 2.11 presents the percent of drivers who were involved in multiple crashes, near-crashes, and critical incidents. Note that close to 50 percent of primary drivers had more than 50 incidents over the course of the study (about one per week) and about 15 percent had more than 150 (about three per week).

Table 2.10. Number and percentage of drivers involved in at least one of the various event types.

Event Type	Number of Drivers	Percentage of Drivers
At least 1 Crash	38	35.5%
At least 1 Near-crash	89	83.2%
At least 1 Incident	99	92.5%

Table 2.11. Number and percentage of drivers involved in multiple events.

Number of Crashes	Percentage of Drivers	Number of Near-crashes	Percentage of Drivers	Number of Incidents	Percentage of Drivers
0	64.5%	0	16.8%	0	7.5%
1	21.5%	1	7.5%	1-5	9.3%
2	6.5%	2-4	27.1%	6-10	3.7%
3	3.7%	5-8	27.1%	11-15	0.9%
4	3.7%	9-12	3.7%	16-20	3.7%
More than 4	0.0%	13-24	13.1%	21-25	5.6%
		25-50	2.8%	26-30	4.7%
		More than 50	1.9%	31-40	8.4%
				41-50	7.5%
				51-100	16.8%
				101-150	16.8%
				151-200	11.2%
				More than 200	3.7%

In viewing the above tables, it becomes clear that some participants might make an outsized contribution to the frequency and rate of events. Further exploration of the matter revealed four participants who might be considered to be outliers when their data is considered on a rate per mile traveled basis. Event rates were calculated for all participants based on event type divided by miles traveled. The rank percentile was then calculated for each event type for each participant. In order to be considered an outlier, a participant had to been > 95th percentile in two of three severity categories and > 90th percentile on the third severity category. Descriptive data for the four participants meeting these criteria are found in Table 2.12. Note that three of the four participants appear to have extremely low miles traveled; for drivers 124, 308, and 311, the miles are an accurate reflection of miles driven. For driver 204, however, there were outages with the data collection system that resulted in apparently low miles traveled. However, the events shown here for driver 204 happened during the miles recorded, so the event rates shown are accurate. Note that the outlier group includes two males and two females, and that four age groups are represented. Driver 311 was one of only 4 female drivers in the 35–to-44 age group, so she might be expected to have a larger influence when age and gender rate calculations are conducted in ensuing chapters of this report. The remaining participants only made up 8 to 11 percent of their respective age and gender categories. The decision was made to include events from these outliers in the remaining sections of this report, with added footnote reminders regarding driver 311 when the 35–to-44 female rates seem unusually high.

Table 2.12. Description of four outlier participants in terms of crash, near-crash, and incidents rates per miles traveled.

Participant #	124	204	308	311
Age Group	21-24	25-34	18-20	35-44
Gender	M	M	F	F
Percentage of age/gender group	10%	8%	11%	25%
Miles Data Recorded	5,241	2,603	4,131	19,833
Incidents	103	60	171	456
Incident rate/mile	0.020	0.023	0.041	0.023
Incident percentile rank	97%	99%	100%	98%
Near-Crashes	7	4	19	56
Near-crash rate/mile	0.001	0.002	0.005	0.003
Near-crash percentile rank	90%	92%	100%	98%
Crashes	3	1	1	4
Crash rate/mile	0.0006	0.0004	0.0002	0.0002
Crash percentile rank	100%	98%	95%	91%

CHAPTER 3: DRIVER AND VEHICLE DEMOGRAPHICS

AGE AND GENDER

Recall that the project goal called for a distribution of driver age and gender, which would slant the study slightly toward the male (60%) and younger (60 percent younger than age 25) end of the spectrum. The original recruiting goals are shown below:

- Age 18-20 years: drivers = 18 males and 12 females.
- Age 21-24 years: drivers = 18 males and 12 females.
- Age 25-34 years: drivers = 6 males and 4 females.
- Age 35-44 years: drivers = 6 males and 4 females.
- Age 45-54 years: drivers = 6 males and 4 females.
- Age 55-64 years: drivers = 6 males and 4 females.

As shown in Table 3.1, the project was successful in achieving the gender distribution goal (60 percent male and 40 percent female). However, the age group recruiting goals were not met. Only 34 percent of participants were under age 25, as opposed to the goal of 60 percent. This was primarily due to the difficulty in trying to recruit participants who drove many miles per year (primarily by commuting). Commuters tend to be older, and younger people tend not to drive as many miles. Those younger participants who were recruited were typically college students who commuted to campus from some distance away. Table 3.2 provides a direct comparison between recruiting goals and achieved distributions for each age and gender grouping. The final distribution did have an advantage, however. As shown in Figure 3.1, the final age distribution was very balanced across the various age groups. As stated previously, some family members and friends would occasionally drive the instrumented vehicles, therefore, data were collected on 148 additional drivers for whom demographic data are not available.

Table 3.1. Participant age and gender distributions.

	Gender		
Age Bins N % of total	Female	Male	Grand Total
18-20	9 8.3%	7 6.4%	16 14.7%
21-24	11 10.1%	10 9.2%	21 19.3%
25-34	7 6.4%	12 11.0%	19 17.4%
35-44	4 3.7%	16 14.7%	20 18.3%
45-54	7 6.4%	13 11.9%	20 18.3%
55+	5 4.6%	8 7.3%	13 11.9%
Total N Total Percent	43 39.4%	66 60.6%	109 100.0%

Table 3.2. Comparison of age and gender distribution to project goals.

Age Bin	Male Goal	Male Actual	Female Goal	Female Actual	Overall Goal	Overall Actual
18-20	18	7	12	9	30	16
21-24	18	10	12	11	30	21
25-34	6	12	4	7	10	19
35-44	6	16	4	4	10	20
45-54	6	13	4	7	10	20
55-64	6	8	4	5	10	13
Total	**60**	**66**	**40**	**43**	**100**	**109**
Total Percent	**60%**	**60.6%**	**40%**	**39.4%**	**NA**	**NA**

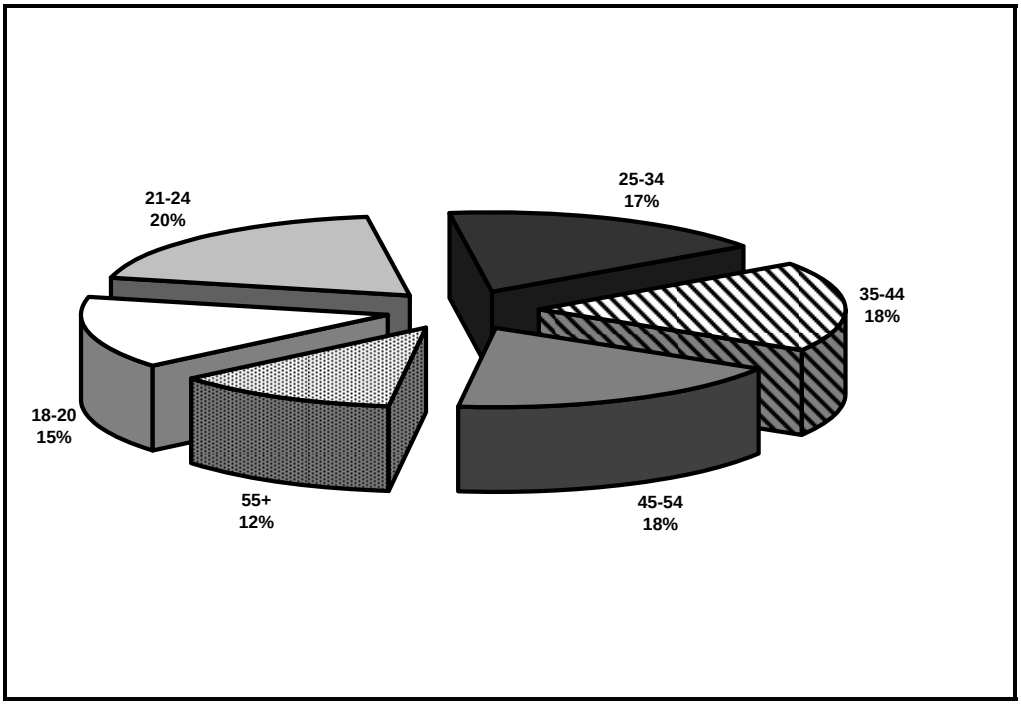


Figure 3.1. Distribution of participant age.

Self-Reported Years of Driving Experience

As seen in Figure 3.2, participants reported a wide variety of years of driving experience. The most experienced group reported greater than 50 years of driving experience (3% of participants), while the largest group was in the 5-to-9 years of experience range (26% of participants). As might be expected, there was a close correlation between driver age and self-reported years of experience, and this relationship is discussed in the next section.

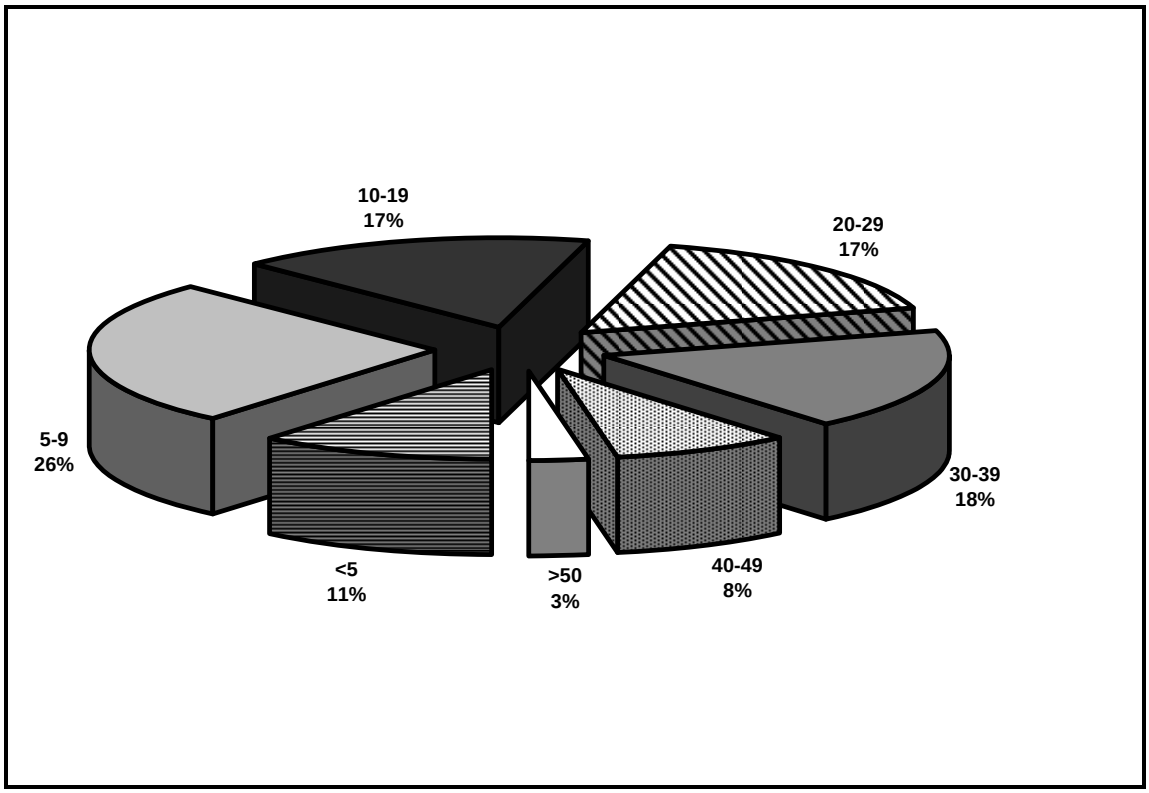


Figure 3.2. Distribution of participant self-reported years of driving experience.

Driver Age as Compared to Self-Reported Years of Driving Experience

In general, self-reported years of driving experience seemed to agree with driver age (Figure 3.3). Note that points above the line represent cases in which the driver began before the average age, and points below the line represent cases in which the driver began after the average age. There are a few data outliers below the line which indicate that 7 or 8 participants began driving three to 5 years after most of their peers, while the one data outlier above the line indicates a driver who self-reported that he/she began driving at age 6. Most likely, this was a mistake on the part of the participant, who probably meant to report 41 years of driving experience rather than 51 years. However, there are cases of people who begin driving at such a young age, especially if they grew up on farms and were required to drive at an early age to help out with the farm work. Females reported beginning to drive at an average age of 17.1 years (SD = 3.51 years), while males began driving at an average age of 16.2 years (SD = 1.94 years). Overall, this pool of drivers began driving at an average age of 16.6 years (SD = 2.69 years).

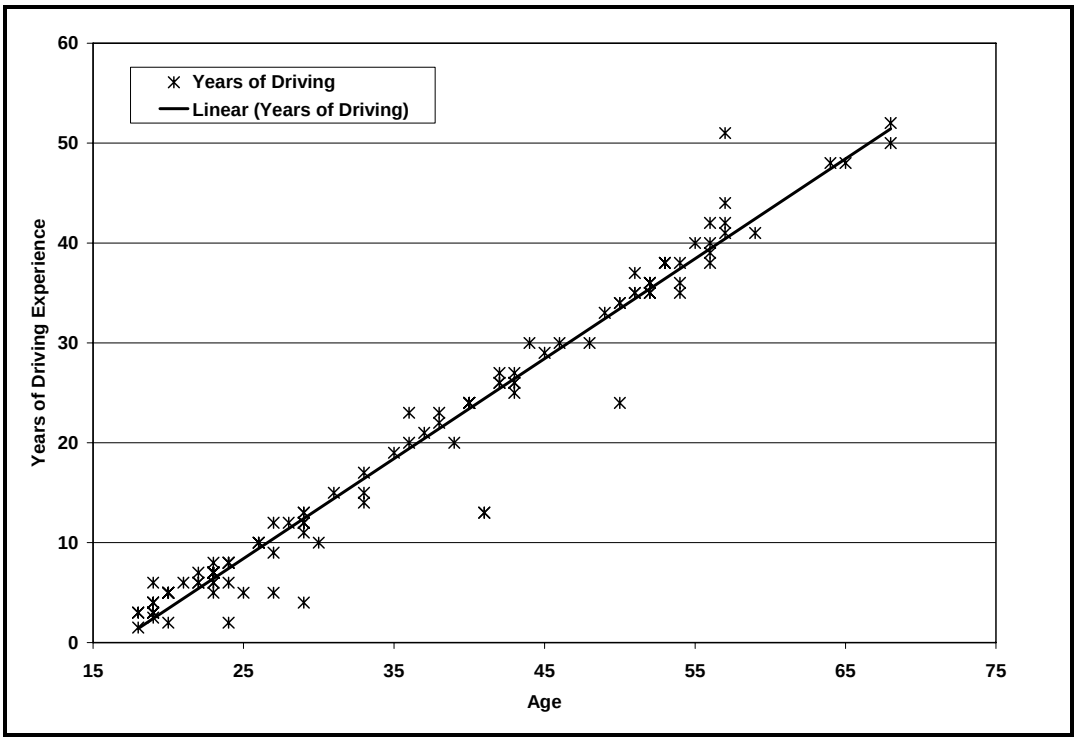


Figure 3.3. Distribution of participants across self-reported years of driving experience bins.

Ethnicity

As shown in Figure 3.4, the pool of participants was largely Caucasian (78% versus 32% non-Caucasian). Table 3.3 shows a comparison of the ethnic makeup of the Northern Virginia area with the participant pool. These data were obtained from a document called "Minority Issues Plan" issued by George Mason University in 2000 (http://www.gmu.edu/departments/provost/accredit/Final%20MINORITY%20ISSUES%20PLAN.doc). Note that the makeup of the participant pool was a fairly close match with the population base of the northern Virginia area, even though no special attempt was made to recruit based on ethnicity.

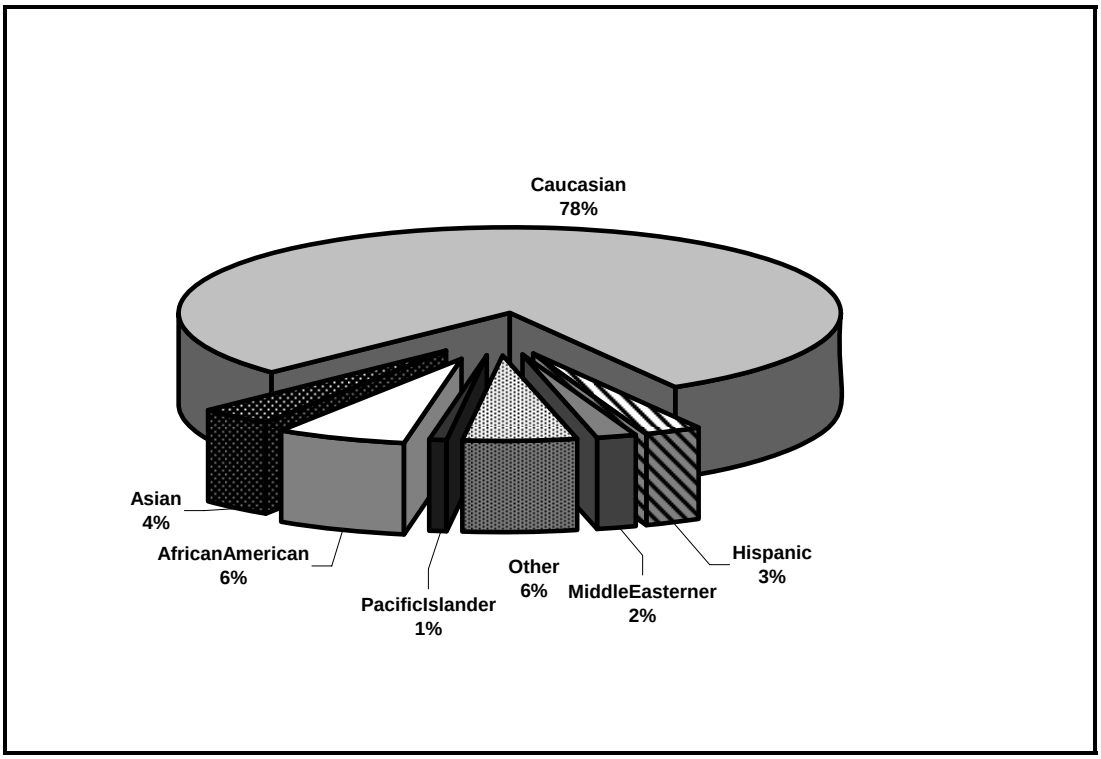


Figure 3.4. Distribution of participants across self-reported ethnic groups.

Table 3.3. Ethnic background of northern Virginia residents and 100-Car Study participants.

Ethnic Group	NOVA Ethnic Makeup	100-Car Participant Pool
White	76.8%	78.9%
African American	9.4%	6.4%
Asian American	6.5%	3.7%
Hispanic American	6.9%	2.8%
Native American	0.2%	0.0%
Other	1.0%	8.2%

Occupation

During the initial screening, participants were asked about their occupations. Their self-reported answers were then placed into related categories as shown in Table 3.4 and Figure 3.5. The categories in Table 3.4 are presented in alphabetical order. The greatest percentage of participants was in the technical field (engineers, drafting, etc. at 18.3% overall), while there were very few who reported being involved in food service, religious, or retired/unemployed (1.8% each). Large differences were noted in the occupations of the participants between genders. For example, of the female participants, 20.9 percent were students, while only 7.6 percent of male participants were students. The situation was almost reversed for the technical category, with 25.8 percent of males in this category and only 7.0 percent of females. Other categories with large differences in gender representation included education (11.6% of females

versus 4.5% of males), legal/military/government (2.3% of females versus 9.1% of males), medical (9.3% of females versus 1.5% of males), and retail/real estate (11.6% of females versus 19.7% of males).

Table 3.4. Occupation categories of participants by gender.

Gender Occupation Category	Female	Male	Grand Total
Education	11.6%	4.5%	7.3%
Financial	7.0%	4.5%	5.5%
Food Service	2.3%	1.5%	1.8%
Legal/Military/Government	2.3%	9.1%	6.4%
Management/Administrative	16.3%	12.1%	13.8%
Medical	9.3%	1.5%	4.6%
Religious	2.3%	1.5%	1.8%
Retail/Real Estate	11.6%	19.7%	16.5%
Retired/Unemployed	2.3%	1.5%	1.8%
Self-employed/Homemaker	4.7%	7.6%	6.4%
Student	20.9%	7.6%	12.8%
Technical	7.0%	25.8%	18.3%
Transportation	2.3%	3.0%	2.8%
Grand Total	100.0%	100.0%	100.0%

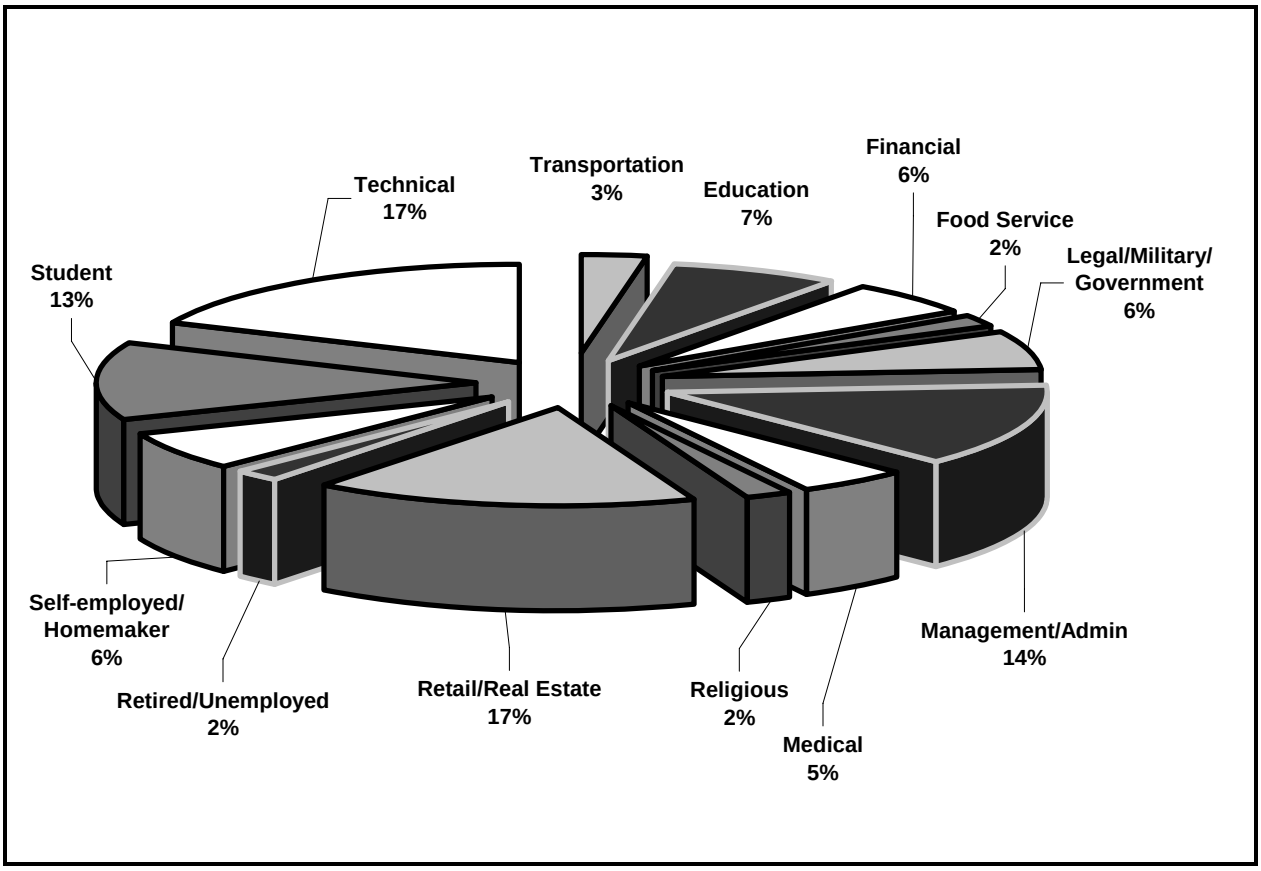


Figure 3.5. Distribution of participants across job categories.

Education

Overall, the participant pool was highly educated (Figure 3.6), with every participant having graduated from high school, and only 2 percent of participants not having at least some college. In fact, 60 percent of the participants reported having at least a 4-year degree, and 19 percent reported having a master's degree, professional degree, or doctoral degree. This is despite the fact that 13 percent of participants were students (presumably working on 4-year degrees) during the time they participated in the study. The overall educational level attained by the participants is probably due to the fact that an attempt was made to recruit automobile commuters. One would expect that automobile commuters would make a fairly good living. They would typically be able to afford housing in the suburbs, a reliable vehicle, and have enough money to purchase fuel for these vehicles. The assumption would then be that people who have attained a higher level of education would obtain jobs that pay adequately to support these items. One would expect that those with lower levels of education might have lower paying jobs, might live closer to work, and if they commuted, they might be more likely to do so via public transportation than by automobile.

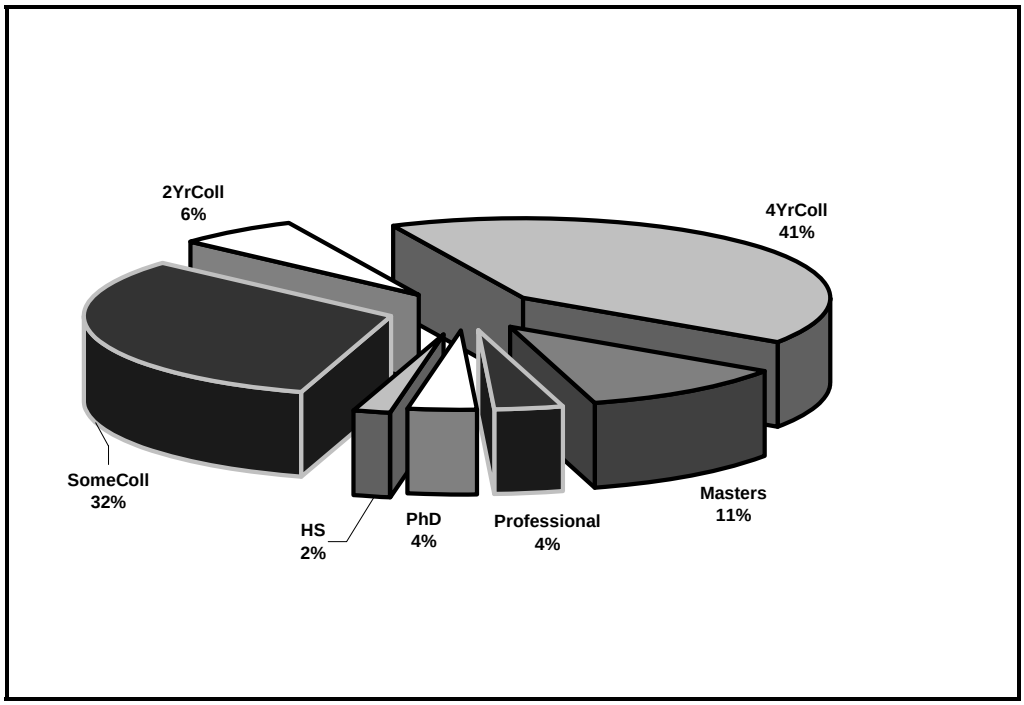


Figure 3.6. Distribution of participants with regard to years and type of education.

Self-Reported Violations

Participants were asked to self-report the number of violations they had received in the past 5 years. The 109 participants reported 163 violations during this time frame, for an average of approximately 0.7 violations per participant in 5 years (0.13 violations/year/participant). As shown in Figure 3.7, the most common category was speeding (63% of violations) and the second most common category was red light violations and stop sign/traffic sign violations (12% each). These three categories accounted for 87 percent of violations reported, with the other 13 percent split among four lesser categories. One participant reported 16 violations, all speeding, while the next highest number for a single individual was 9 violations (all red light running violations). There were 37 participants (34%) who reported no violations during the past 5 years, and 2 participants who did not answer the question.

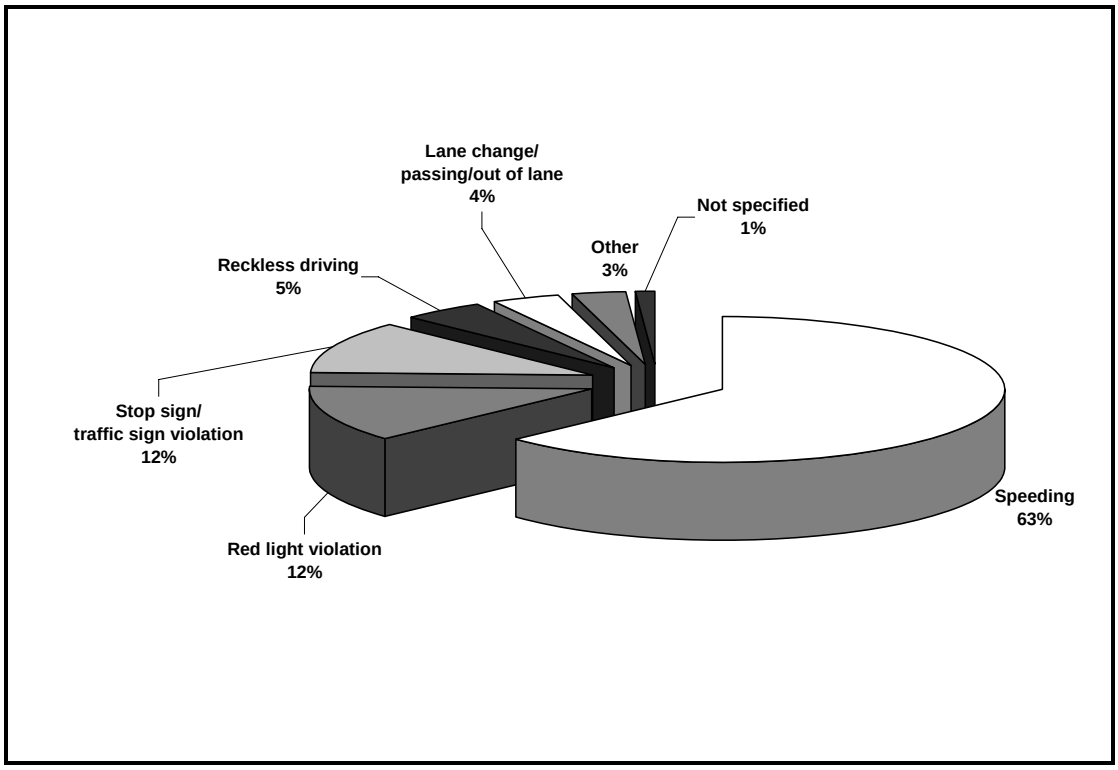


Figure 3.7. Distribution of type of self-reported traffic violations in the past 5 years.

Gender differences were noted in the distribution of the number of self-reported violations (Figure 3.8). About 25 percent of females reported no violations in the past 5 years, as compared to about 40 percent of males. Over half of females reported having 1 or 2 violations in this time frame as compared to about 40 percent for males. Males and females reported having three or more violations at about the same rate (around 20%). Since these violations were self-reported, there is no way to ensure the accuracy of these figures.

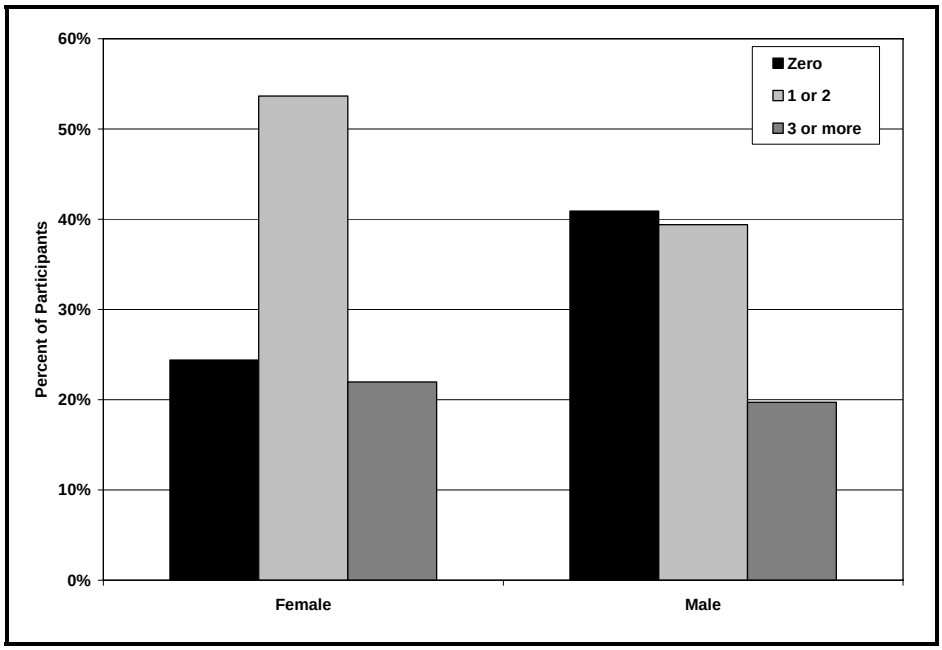


Figure 3.8. Percentage of participants by gender for number of self-reported traffic violations in the past 5 years.

Some interesting age trends were also noted with regard to self-reported violations. There was a distinct split between those younger than age 35 and those 35 or older, as shown in Figure 3.9. For each of the three age groups younger than 35, fewer than 20 percent reported having had 0 violations in the past 5 years. In contrast, the three age groups aged 35 or over had 45 percent or more of participants reporting no violations during this time span. The three or more violations category peaked in the 21- to 24-year-old age group, at nearly 50 percent. Note that for a 21-year-old, the past 5 years encompasses their entire driving history, assuming that they began driving at age 16. The three violations or more category was lowest between the ages of 35-55, at 5 percent or less.

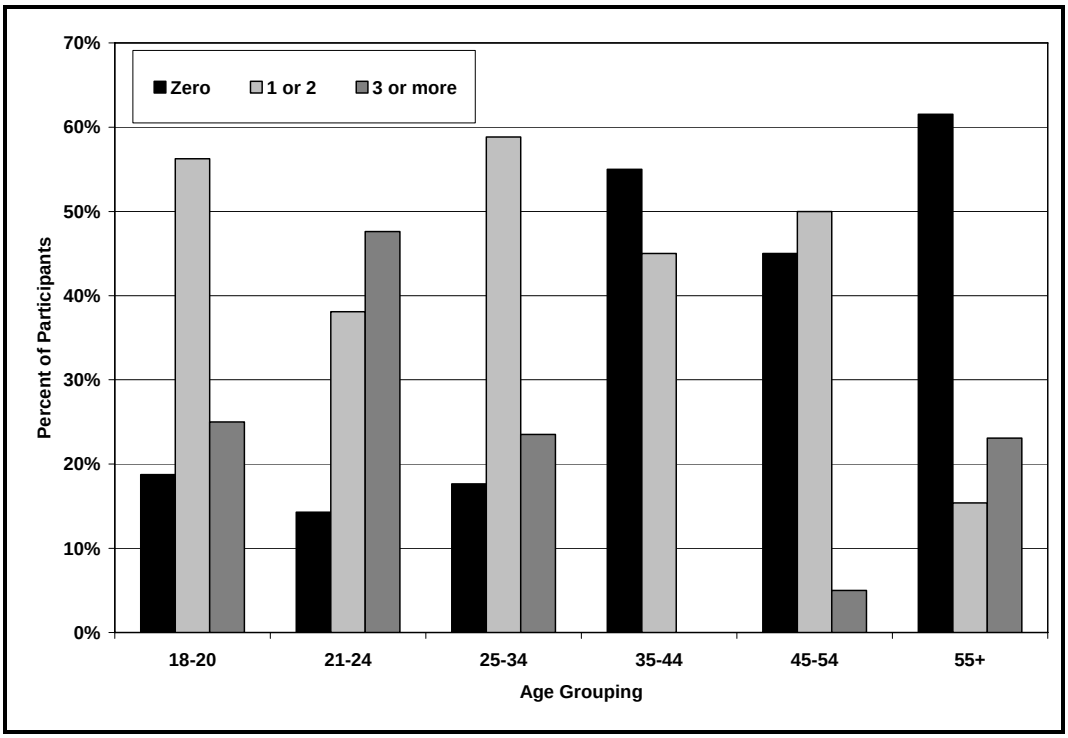


Figure 3.9. Percentage of participants in each age bin for number of self-reported traffic violations in the past 5 years.

Self-Reported Crashes

A similar question asked participants to report the number of traffic crashes they had been involved in over the past 10 years. Overall, 35 percent of drivers reported no crashes over this time span, 50 percent reported one or two crashes, 13 percent reported 3 or more crashes, and 2 percent did not answer the question. Unlike for violations, there was close agreement among the genders for this question (Figure 3.10).

There were also age trends for this question, but the differences were not as pronounced as for the violations question (Figure 3.11). For the zero crashes category, the peak age group was 45- to 54-year olds at 60 percent and the lowest age group was 25- to 34-year olds, with only 10.5 percent of this age group reporting no crashes over the past 10 years. Of the 21- to 24-year-olds, 29 percent reported having been in 3 or more crashes over the past 10 years. In all cases, this likely represents their complete driving history, even for those who might have started driving at a younger age. In contrast, only 5 percent of 45- to 54-year-olds reported 3 or more crashes over this time frame.

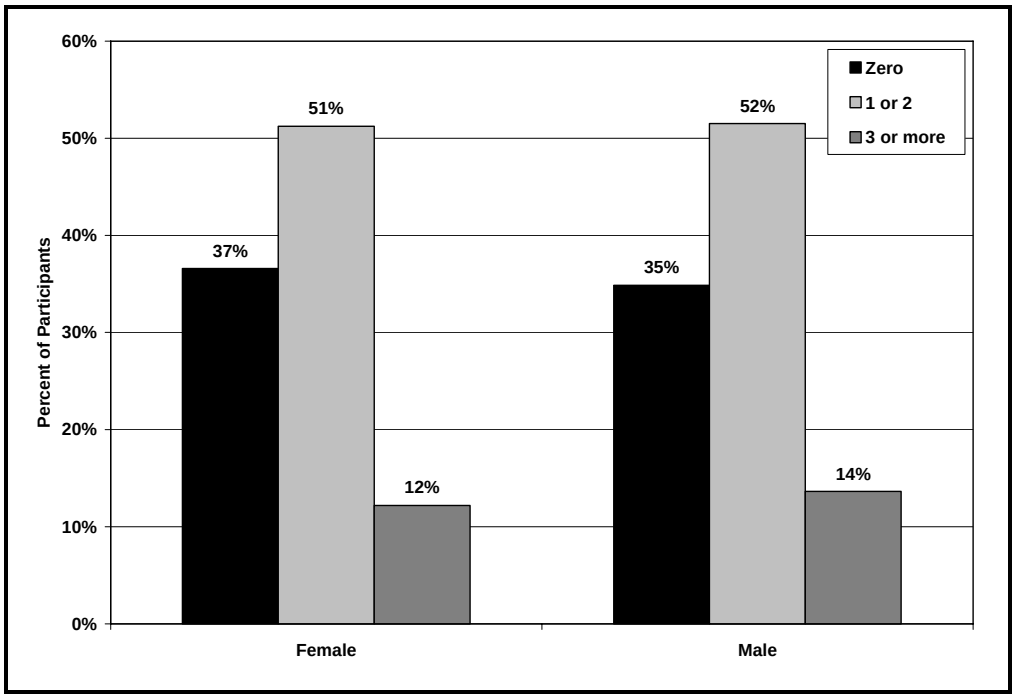


Figure 3.10. Percentage of participants by gender for number of self-reported traffic crashes in the past 10 years.

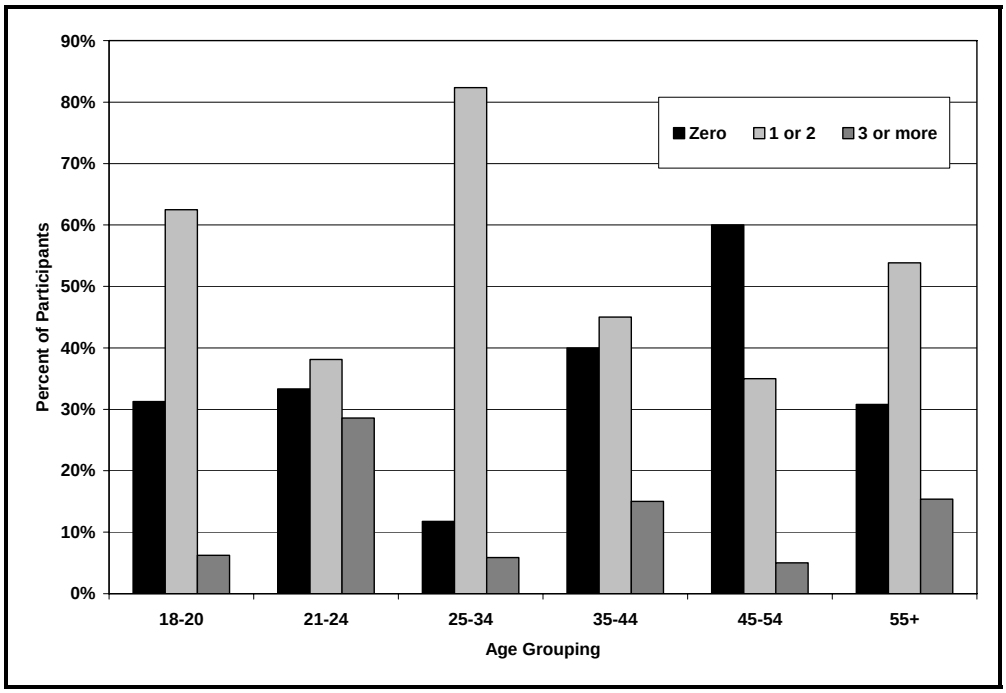


Figure 3.11. Percentage of participants in each age bin for number of self-reported traffic crashes in the past 10 years.

Self-Reported Annual Mileage

Participants were asked how many miles they typically drove in a 12 month period, in keeping with the desire to recruit high mileage commuters for the project (Figure 3.12). Over three-fourths of participants reported driving more than 15,000 miles per year, but these numbers were shown to be inflated based on mileage driven during the study period. Table 3.5 presents the number of drivers for each of several actual mileage bins. As can be seen, in reality only 30 percent of participants drove more than 15,000 miles during the course of the study, and one-fourth of participants drove 9,000 miles or less during the study. Nevertheless, the project goals in terms of numbers of crashes, near-crashes, and crash-relevant conflicts were met, so this inflation of miles driven did not turn out to be important to the outcome of the study.

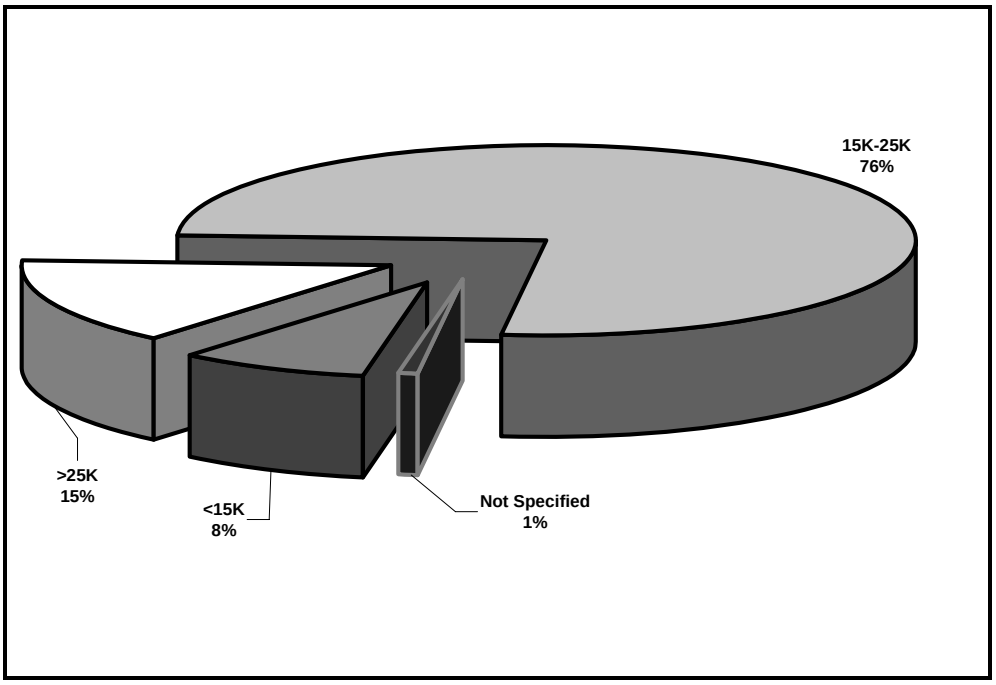


Figure 3.12. Distribution of participants across self-reported annual mileage categories.

Table 3.5. Actual miles driven during the study.

Actual miles driven	Number of participants	Percentage of Participants
0-9,000	29	26.6%
9,001-12,000	22	20.2%
12,001-15,000	26	23.9%
15,001-18,000	11	10.1%
18,001-21,000	8	7.3%
More than 21,000	13	11.9%

Vehicle Demographics

Overall, 78 vehicles were privately owned and 22 were leased vehicles loaned to the participants in return for their participation in the study. As can be seen in Figure 3.13, there was quite a large difference in the distribution of leased versus privately owned vehicles with regard to age groups. Younger drivers were more likely to be motivated to participate in the study by the prospect of having a vehicle to drive with no car payments due, while older participants were less likely to want to give up the familiarity of their own vehicles. No drivers age 30 or over used a leased vehicle, while over 60 percent of the 18- to 20-year-olds drove leased vehicles, falling to just over 25 percent for participants in the 25-34 year age category.

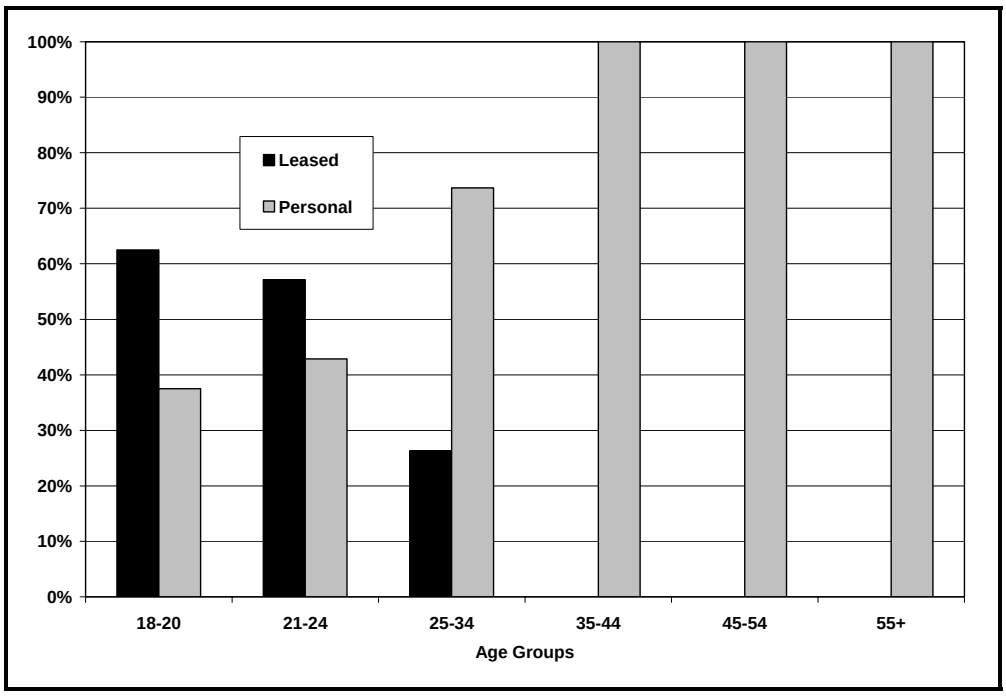


Figure 3.13. Percentage of participants for each age group driving personal versus leased vehicles for the study.

An attempt was made to have the chosen vehicle type makes and models evenly represented in the study. As can be seen in Figure 3.14, this goal was achieved, with a 12-20 percent share for each make/model combination. Likewise, the three manufacturers were fairly evenly represented, with 38 percent Chevrolet, 27 percent Ford, and 35 percent Toyota.

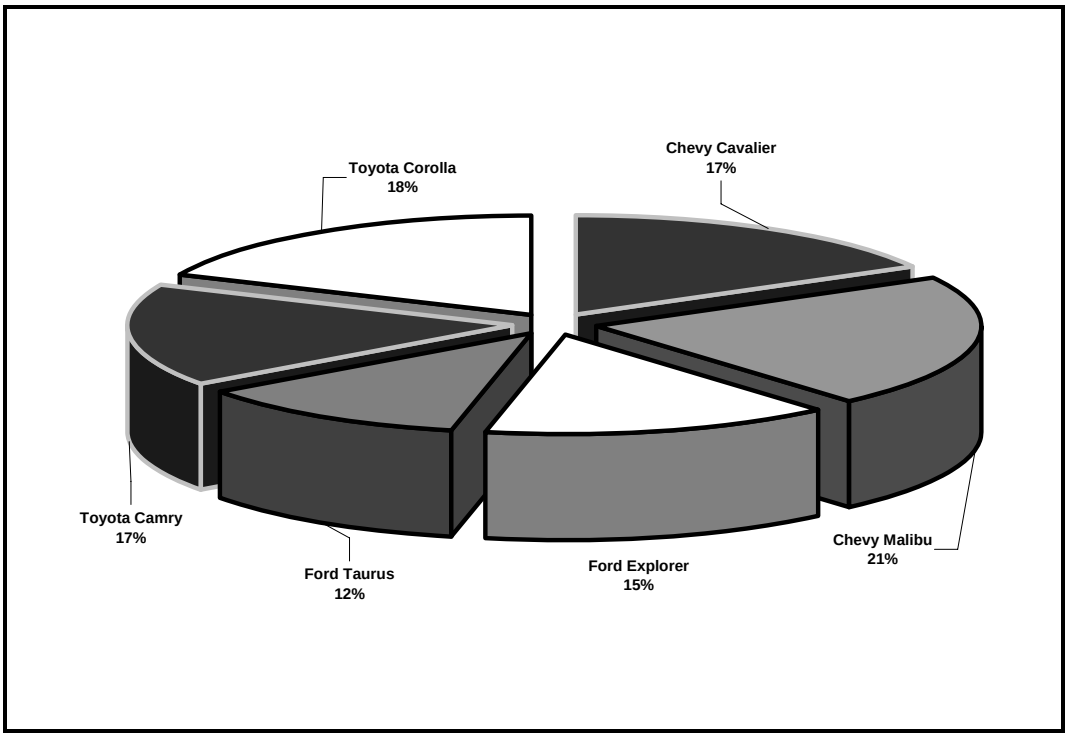


Figure 3.14. Distribution of participants across vehicle makes/models driven during the study.

This page intentionally left blank.

CHAPTER 4: LESSONS LEARNED IN CONDUCTING THE 100-CAR STUDY

BACKGROUND

The logistics and theoretical implications of conducting a field experiment of the scope and complexity of the 100-Car Study proved to be challenging from the beginning to the end of the project. Essentially, every aspect of the project differed from the norm either because of the number of vehicles, the amount of data gathered, or because new software reduction techniques had to be developed to take full advantage of the data gathered. When challenges arose, solutions were promptly created to overcome them with minimal impact on the dataset. In addition to the current and future value of the dataset, there was a substantial amount of organizational knowledge created as part of this effort, which can best be presented as a series of *lessons learned.*

These *lessons learned* should serve as recommendations to ease the performance and management of any future naturalistic driving studies of similar or larger scope. The remainder of this chapter discusses the challenges encountered in the study as well as recommendations to address them in future efforts. This discussion is framed within the following broad categories:

1) Subject Recruitment and Compliance
2) DAS Installation
3) Hardware and Software Maintenance
4) Data Downloading
5) Data Reduction
6) Data Analysis
7) Other Logistics

SUBJECT RECRUITMENT AND COMPLIANCE

The recruitment of subjects to participate in the study was challenging due to a combination of several screening factors. One factor was the selection of only six different makes and models of vehicles to be included in the study. During Task 8 of Phase I, the vehicle types were chosen to limit the number of customized bracket types that had to be created. Although VTTI purposefully chose vehicle makes and models that were popular in the northern Virginia area, this selection still narrowed the number of drivers who could have their private vehicles instrumented. The participant pool was further reduced due to the driver's ages that were needed, the requirement that a high number of miles were typically driven, and the limit on the targeted geographical area. The reduced participant pool made the participant recruitment process somewhat difficult.

Furthermore, the vehicle models selected for inclusion in the study were typically not driven by younger participants. Even when the younger drivers drove the particular type of vehicle, the vehicle was typically an older model year with a different body style requiring the creation of different mounting brackets. Rather than continue to create brackets that would only be used for a couple of vehicles, leased vehicles were used by a large portion of this study group. This option represented the most efficient way of incorporating the younger driving population within the study. Any future studies should be aware of the importance of a large and diverse subject

pool, and avoid geographical areas with relatively small populations, unless the experimenters are prepared to customize the DAS for a large number of vehicles makes and models.

Driver attrition and/or removal of the driver from the study were also important aspects of subject management and required additional drivers to be recruited throughout the study. Apart from a small number of drivers who exercised their right to withdraw from the experiment, four particular cases are indicative of the complexity of the issues that can arise when participants are tracked for long periods of time:

- A driver moved away from the study area.
- A driver was arrested during the course of the study and could no longer participate.
- A driver of a leased vehicle was in three crashes and Virginia Tech Office of Risk Management no longer wanted to cover the insurance.
- A driver of a private vehicle that experienced a catastrophic mechanical failure and it was not economically feasible to repair it, therefore, the driver could no longer participate.

Future studies should always have a small number of "reserve" participants who can be called on relatively short notice to replace any drivers removed from the study. Thus, participant recruitment and initial screening should continue beyond the placement of the desired number of vehicles on the road.

Subject compliance issues were also present in the study. Despite numerous efforts to explain the study protocol to drivers and to relay the importance of their compliance, several drivers chose not to do so completely. Some interesting examples include:

- A driver of a leased vehicle loaned the vehicle to an unlicensed driver, thus violating the study protocol.
- A few drivers would not come in for the final debriefing, even though they would be paid $150.00 for less than an hour of their time.
- Some drivers would not report damage to the leased vehicles, even though failure to file a police report required payment of a deductible that they would not otherwise have to pay.

These examples point to the importance of the person or persons who are in direct contact with the participants and who serve as the interface between the participants and the organization performing the study. These employees should be well trained in working with participants and with the resolution of the unique issues that are likely to arise in a study of this length and magnitude.

DAS INSTALLATION

The 100-Car Data Acquisition System was highly capable and complex, yet had to be installed in privately owned vehicles without any permanent vehicle modifications. To achieve this, VTTI engineers developed customized brackets to utilize existing mounting holes in the frame of the vehicle. However, in some cases the tolerances for the placement of these mounting holes were larger than expected. Therefore, brackets that should fit a particular vehicle sometimes did not

and required a certain amount of customization. Prevention of this problem requires the introduction, within the bracket system, of a certain amount of adjustability.

Installation quality control issues also arose due to the hiring of a subcontractor to assist with installations. If a subcontractor is hired to perform installation, maintenance, or repairs, the selection process should carefully consider the capabilities of the contractor, their willingness to receive specialized training, and their typical level of customer service (i.e., timeliness of work, craftsmanship, politeness, and attention to detail). In addition, strict guidelines as to who will be responsible for repair and payment for an installation problem, detailed instructions for installations, and explicit expectations for the installation timeframe are all critical. It was found that if participants were asked about their experiences during the installation, and subsequent feedback was provided to the garage, the subcontractor performance levels improved. Regardless of the care with which the installer is selected, or even if the systems are installed by in-house personnel, these surveys should continue for any future studies. It might also be useful to institute random inspections of recently instrumented vehicles to catch any systematic problems with the installation that require further training or information to the installer. These inspections may be useful even when the installations are performed by in-house personnel.

Some sensor installation issues also existed, especially related to crash survivability. The VORAD units and brackets, installed in front of the bumpers, were destroyed or damaged in most of the crashes, even when the crash was relatively minor. While other installation options could be explored to place the radar unit behind the bumper or grille, the particular placement of the radar was a cost effective solution for this study.

With regard to the radar, standard license plates could not be used since the radar could not "see" through the metal. VTTI staff worked with the Virginia Department of Motor Vehicles (DMV) to have special plastic plates manufactured. The downside of the plastic plates was that the plates were fragile and had to be replaced on multiple occasions for many different cars.

HARDWARE AND SOFTWARE MAINTENANCE

The single most important lesson learned regarding system maintenance for the 100-Car Study was to have a maintenance person permanently located within the northern Virginia area near the study vehicles. Initially, hardware and software maintenance was performed periodically by in-house personnel, who would travel to the northern Virginia area to perform repairs on the vehicles that needed them. These employees took replacement parts and tools with them, but had to perform the repairs on the road, wherever the vehicle was parked at the time that it was intercepted. Having personnel in the area was helpful in that they were able to respond to problems more quickly, had more familiarity with the roads on the area, and had a permanent space in which to make repairs, reducing the time that had to be spent with the car and the possibility of inconveniencing participants.

Equipment repair and adjustment times were also reduced by allowing data downloaders to perform minor work on the cars. This work was typically performed during the time that data were being downloaded. This approach reduced the work load of repair personnel, whose effort could then be focused on more complex repairs that required higher levels of technical expertise.

Hardware

The 100-Car Study DAS had a unique remote tracking capability that allowed study personnel to determine, based on GPS coordinates, the location of a vehicle. This functionality was essential when data downloaders had to locate the vehicles. The system was also able to transmit limited amounts of data from the vehicle that could be used for fault detection. In some cases, the system also allowed for the remote completion of small repairs to the system, especially those that involved resetting particular pieces of equipment. This capability should be maintained or expanded in future systems, if feasible.

Repairing the system posed several substantial problems. The design of the DAS as a closed system had the advantage of making the DAS fairly unobtrusive and reduced the possibility of equipment tampering. However, it also posed problems for repair personnel, since removing the DAS was not a simple task. Unfortunately, all repairs involving a component residing within the box required that the box be removed. Future iterations of the system should, despite any possible improvements in the ruggedness and robustness of the components, provide for an easier DAS removal process. This could be achieved via alternative mounting approaches or by the use of special mounting tools available only to system maintenance personnel. Another option would be to modularize system components so that they can be added or removed through access doors to the DAS. This would also reduce the overall time required for a repair.

Similarly, the internal layout of the system was not optimal in allowing quick repairs. Replacement of the boards located lower within the system (e.g., the motherboard, video board, and quad splitter) required complete DAS disassembly. Future designs will likely integrate more of these functions within a single board, thereby reducing the number of separate components. They will also streamline the DAS layout to allow for faster repairs.

Hard drive management was also a substantial maintenance challenge. While the hard drives were fairly robust, there were a number of failures resulting in lost data. As hard drive technology improves, their failure rate should decrease, but losing data will still be a possibility. Alternative data storage methods not involving moving parts could be explored, along with the possibility of redundant data storage. Alternatively, the hard drive could be placed so that it is accessible to repair personnel without removal of the DAS. However, the downtime due to failures of this type, and consequently the amount of data lost, was relatively small, as discussed in Chapter 13, Goal 9.

Despite substantial efforts to prevent it, several data acquisitions systems drained the batteries of the cars in which they were installed. Safeguards to prevent this problem included the provision of an internal battery backup system that could be used to operate the system while the vehicle was turned off (e.g., when data were being downloaded), the inclusion of a software switch that turned the DAS off if the voltage of the car battery dropped below 11 V, and the inclusion of a "suicide" feature that automatically shut the DAS down when the vehicle was turned off (except for data download purposes). However, these safeguards failed in some cases. For example, the system would keep running after the car was turned off, which occurred when the operating system was not working properly. In those instances, the system kept resetting the CPU to restart the operating system, thereby draining the vehicle battery in the process. These incidents inconvenienced the participants, and should be avoided to the largest extent possible in future

efforts. Better sensors and more robust system shutdown algorithms can be created to address the majority of these issues, and should be implemented in the future.

Another minor issue concerned system time, which could be off by several hours after a few months of data collection. It is not known at this time whether this particular problem was due to the operating system or to the motherboard. This problem should be addressed and tested in future DAS designs. A more severe version of the problem occurred when some motherboard jumpers shifted off of their intended positions, resetting the system time. Hard-wiring the two terminals solved the problem permanently. In the future, hard-wiring any potential moving parts within the motherboard should be performed prior to DAS deployment.

A final hardware aspect had to do with the proper installation of all sensing equipment. When necessary, installation verification tools should be created and used to ensure consistency between vehicles. For example, between-vehicle consistency of radar alignment could be improved. Radar units were not checked for perfect alignment or orientation, although the errors were less than 3 deg. Since azimuth of target was used as an exclusion criterion, it is possible that headway trigger performance (which used the radar sensor) could have been improved by ensuring that the radar units were perfectly aligned.

Software

The main software-related challenge was to make the DAS work on a Microsoft Windows™ 98 operating system. This operating system is not designed as a real-time operating system, thereby creating some issues with data synchronization, output, and storage. The operating system also had a relatively slow data transfer rate for networking operations. Given that the data download process was performed though a network link, it tended to be a lengthy process, which in turn could drain the car battery, as discussed in the previous section. The Windows™ 98 operating system was also susceptible to power failures and/or system crashes, and in some circumstances, resulted in disk boot errors. Finally, the operating system also allowed the hard drives, on occasion, to continue to collect data until all of the disk space was used. This prevented the operating system from booting and, in some cases, caused the corruption of the data within the hard drive. These catastrophic failures, while problematic, were not a frequent occurrence.

All of these software problems with the operating system suggest the use of a different operating system for future systems. This operating system will likely be Linux-based, since these operating systems solve many of the problems that were evident when using Windows™ 98. Linux-based DASs are already in operation and will likely be used in any future experiments of this type.

Another software-related challenge was the number of DAS software versions released, especially during the initial stages of data collection. Despite numerous efforts to debug the system before initial use, minor modifications were necessary. These modifications were completed as needed and downloaded into the vehicles gradually, but this implied that many different versions of the data collection program (XCAR) were operating in the data acquisition vehicles at the same time. Understandably, this caused confusion as to which was the latest or most appropriate version of the software. Several improvements to this process can be suggested. The most obvious is to increase the test time before systems are deployed. Future

systems should benefit greatly from the expertise gained in this project so that the number of software bugs in future data collection systems will be greatly reduced. Another improvement would be to schedule software releases instead of sending out new versions meant to fix a single bug. While this would not be recommended for a major bug (e.g., a bug which prevents data collection), it might be possible for relatively minor fixes, therefore allowing more time to ensure that all systems are updated before a new software version becomes available.

DATA DOWNLOADING

The main difficulty with data downloading consisted of gaining access to vehicles. Some participants did not fully cooperate with the data downloaders when it was time to download. In most instances, downloaders did not need access to the vehicles or direct communication with the driver since the cars could be located remotely and data downloaded unobtrusively. Something that aided in this respect was to obtain the participants' regular schedules in advance. This allowed downloaders to schedule their visits in advance with minimal inconvenience to the drivers (who were unlikely to need the vehicle during that time). However, in situations in which the downloader needed access to the car to fix a sensor (e.g., correct the orientation of a camera), they needed to interact with the driver to obtain access to the car. Some drivers were more cooperative, in this regard, than others.

In a related issue, detailed logs had to be kept of the data downloads for each of the vehicles. This allowed the downloader to prioritize vehicles according to the amount of data not yet downloaded, thereby minimizing the risk of data lost due to a full hard drive. Thus, the decision of when to download which vehicles was not only dependent on a participant's schedule but on the amount of data stored in the vehicle. In addition, downloaders had to work a flexible schedule that allowed them to access some cars in the evening hours when participants were less likely to be using the vehicles.

The data downloading process also requires careful consideration of data security, archiving, and storage issues. Server managers kept detailed logs of the data sent from the northern Virginia location and the data received in Blacksburg. These logs were periodically compared to ensure that no data were missing. In addition, backup copies of the data were maintained in various locations in order to minimize the risk of data loss. Downloader laptops were cleared of data within one day of the data download. The main lesson learned in this respect is that close communication and interaction is needed between data managers and server managers to ensure a smooth and complete data flow.

Finally, the geographical area encompassed must be considered when determining the location and number of downloaders. If the geographical area of future studies is large, then multiple downloader "bases" could be considered to reduce response times.

Note that these issues would not be directly applicable to a large-scale naturalistic study, as the download process in that case will likely be different. For such a large number of vehicles, data download stations would be set up, and participants would return to the station at the end of their time in the study. While this would eliminate a large portion of the problems related to "chasing" cars, some of the data downloading lessons learned as part of this study (e.g., maintaining detailed repair logs) still apply.

DATA REDUCTION

The data reduction process for this study was developed to record epidemiological data, similar to the GES crash database, as well as record data that has typically been collected in other instrumented vehicle studies, thus greatly augmenting both types of data collection. The 5 channels of video were primarily used to record these variables. However, the data reduction software, developed in-house, allowed the data reductionists to access time plots of the various vehicle sensors (i.e., longitudinal deceleration, vehicle speed) and could be used to record certain other variables as well, e.g., a complete list of data reduction variables is located in Appendix B.

Even with driving performance data and video greatly enhancing the data reduction process, many reduction variables still required a judgment call or subjective analysis on the part of the data reductionist. Many steps were taken to ensure inter-rater reliability and reduce subjectivity among the data reductionists for these types of variables. First, a two-week training process was provided for each reductionist to allow them to:

- Learn the data reduction software,
- Practice viewing all 5 channels of video,
- Understand the trade-offs of using the video versus using the driving performance time plots, and
- Work with both the lab manager and other trained reductionists to develop a broad understanding of the types of judgments that needed to be made.

Second, all data reductionists were expected to attend weekly meetings in which questions and issues about various data reduction topics were discussed. Third, the lab manager(s) performed spot-checks of all reductionists' work and provided individual feedback to the reductionists. Reductionists were also required to spend 30 minutes each week spot-checking other reductionists' work and providing feedback/discussions to these reductionists. This step was useful for two reasons: (1) it improved accuracy in the database, and (2) it allowed the reductionists to observe other's work and conduct a comparison to their own work, thereby increasing consistency among all reductionists. Finally, three inter-rater reliability tests were conducted in which the reductionists were all required to validate the same 20 events (per test) and fully reduce two of the twenty events. The test results indicated that there was 88 percent inter-rater reliability for validation of events and 99 percent intra-rater reliability for recording all of the reduction variables. An interesting anecdote is that the inter-rater reliability tests proved to be a very beneficial training tool and will be used from the earliest stages of future data reduction efforts.

Because more information was available to the data reductionists than to the GES analysts who enter information from police-accident reports into the GES database, many of the GES variables were expanded for this study. The GES database is for crashes only, so some of the GES variables were not included in the 100-Car Study database because they were not applicable (e.g., occupant injury, EMS response times). As the reduction process began, the high variability among the events and among the drivers became more apparent. Nevertheless, coding a pre-incident maneuver, precipitating factors, contributing factors, and evasive maneuvers for each event, as well as coding a pre-incident maneuver and evasive maneuver for each vehicle involved and surrounding the event, appeared to adequately capture the pertinent information for

the vast majority of the events. Having the data reductionists write a narrative, or written description of each event, allowed other useful information to be recorded and used for future analyses.

Incorporating 5 video channels were incorporated into the 100-Car Study DAS was done to ensure the capture of as much of the drivers' view surrounding the vehicle as possible (forward view, rear-view, rear-facing passenger window, and outside the driver's window, via the angled face camera), as well as driver behavior (face view and over-the-shoulder view). There are trade-offs associated with these 5 camera views, which include size of video files and resolution of the video. Five channels of video increased the bandwidth of the video data, which forced VTTI engineers to decrease the level of resolution of the video so that storage issues would not become problematic. However, the resolution level provided by the system still allowed eyeglance reduction to be performed. The resolution levels had a higher effect on discriminating objects and obstacles outside the vehicle. Potholes, for example, were very difficult to identify. Street signs (i.e., speed limit signs) were not readable. Objects inside the vehicle were also sometimes difficult to identify in the camera views. Any problems due to resolution were compounded by nighttime hours (in which visibility is lower) and sunlight glare (which "washes-out" the camera). These aspects also made eyeglance reduction much more difficult, although still possible in most cases. While technological advancements in video have already addressed many of these problems, the usefulness of all 5 video channels should be addressed prior to a large-scale study and trade-offs between video resolution and additional channels of video should be weighed carefully.

DATA ANALYSIS

The data analysis process for this amount of data proved to be challenging, time-intensive, and complex. The main lesson is to allow enough time for databases to be created. When variables were derived from the raw data (i.e., vehicle miles traveled), substantial processing time was required (as much as one week of processing time in some cases). Also, different analyses used different subsets of subjects. For example, demographic data were only collected for primary drivers of vehicles; therefore, if age or gender was necessary in the analysis, only 109 subjects' data could be used. If simple frequency calculations were used, then all drivers would be eligible for the dataset. Therefore, 241 subjects would be used in the analysis. These differences have implications for frequency and rate counts for crashes, near-crashes, and incidents.

Given that different analyses required different subsets of data, the decision was made to keep the data centralized, so that only one or two people conducted queries and performed statistical analyses. While this procedure may seem inefficient, having one person in control of producing datasets was imperative to maintain consistency throughout the report. This process only works when enough time is allotted to conduct analyses as the database manager can quickly become a bottleneck in the flow of analysis and report writing.

Even with a single person in charge of the query process, the number of people who worked with the resultant data turned out to be large. This made the analysis revision process much more difficult when errors were found or changes made to the database or the reduction structure. A formal communication structure is suggested in the future so that all relevant personnel are informed of changes. In addition, a log for these changes should be created and maintained by

the data manager(s) so that a record exists of any changes made to the data or database and as well as the reasons for the changes.

Finally, in some instances it would have been useful to observe certain kinematic variables at a higher sampling rate. This capability is suggested for the future, even if it requires that only triggered data are stored (to reduce needed data storage capacity). This is certainly compatible with the trigger based data collection system that is foreseen for any future larger-scale studies.

OTHER LOGISTICS

The use of leased cars (owned by the Motor Pool department at Virginia Tech) was useful in many regards, but proved to be a hindrance in many other ways. Significant problems had to be overcome in order to obtain the vehicles and then change their registrations so the vehicles did not use state plates. This last action was necessary because some participants were aware of the benefits of using a vehicle with state plates (e.g., reserved parking spots at state universities) and misused them. In addition, maintenance for the leased cars, a responsibility of the participants, was sometimes neglected. The Motor Pool at Virginia Tech has rather strict maintenance schedules in place, and the leased cars had to adhere to these schedules. When a vehicle was overdue for maintenance (e.g., an oil change), substantial time and effort had to be invested in either getting the participant to service the car or intercepting the car so that our personnel could service the vehicle.

Another aspect of using leased vehicles was the necessity of keeping a log of miles traveled per month, information required by Virginia Tech's Motor Pool department. Obtaining this information from drivers every month proved to be a very difficult and time consuming task.

These drawbacks should be considered in any future efforts of a similar or larger scale. While there are advantages to leasing cars for the study, there is a large procedural overhead that accompanies the leasing of such vehicles, given that the Motor Pool is a department within an entity in the Commonwealth of Virginia, and thus constrained by governmental protocols.

Another important logistical issue was the coordination with the Virginia Department of Motor Vehicles (DMV) to obtain approval for the plastic plates and the need to re-register participant's vehicles under the plastic plates (and then back to the original plates once the study ended). A related issue was that the plastic plates did not have a retro-reflective coating since this caused distortion of the radar signal. As a precaution, VTTI staff coordinated with the Virginia DMV to have a letter from the DMV commissioner stating that the driver of the vehicle was a participant in a study and the plastic plates were sanctioned by the DMV. Setting up this process was very time consuming, as the decision to approve all of these measures had to be taken by central office personnel at the DMV. Several DMV offices then had to coordinate efforts in order to make the registration process quick and simple. However, once the registration process was established and contacts were made with key people, it went smoothly and registration materials for new plates were obtained within a few days of the original request. In most cases, participants received all the necessary DMV materials the same day that the installation took place.

Finally, important lessons were learned with regard to protecting the confidentiality of the drivers in the study. To protect the drivers in the event of a crash, it was deemed important to obtain a Certificate of Confidentiality from the National Institutes of Mental Health (NIMH). The purpose of this certificate was to prevent the data collected in the study from being subpoenaed so that it could not be used against a subject in court. However, obtaining the certificate imposed a constraint on the study. Specifically, it was an original goal of the study to instrument the vehicles to collect video of the entire cab of the vehicle as well as to collect audio continuously to better understand the effect of passengers on driver distraction. Nonetheless, administrators at NIMH felt that it was important to protect the confidentiality of anyone in the vehicle who could be recorded via video or audio recordings. To have the driver administer and submit informed consent forms (or assent forms for minors) for every person who may get into the vehicle during the course of the year was considered infeasible and inappropriate. Posting a message inside the vehicle telling every person that they were being recorded was thought to have a negative effect on the naturalistic data collection approach with regard to the driver. Therefore, the choice was made to use camera placement and angles that would only collect data on the driver and to only have audio recording active when the driver activated the incident push button. Obviously, from the perspective of understanding the degree to which passengers are creating a distraction in the vehicle, the data collected are not as complete as initially desired.

DISCUSSION

These aspects represent the major issues that had to be addressed throughout the data collection and analysis period. While sizable, they were all addressed satisfactorily, and should not, in the majority of cases, present significant issues in future studies, even if the study is of a larger magnitude. Every study brings new challenges, and perhaps the most important lesson to learn from this substantial effort was that the organizational desire and determination to correctly address issues is alive and well within the organizations that came together to perform this work.

CHAPTER 5: GOAL 1, CLASSIFY AND QUANTIFY CAUSAL FACTORS AND DYNAMIC SCENARIOS INVOLVED IN EACH CONFLICT CATEGORY.

DATA ANALYSIS OVERVIEW

For this research goal, the crashes, near-crashes, and incidents were parsed into the following 18 conflict categories. These conflict categories are found in many crash databases and provide a common, consistent method to stratify the data.

- Conflict with a lead vehicle
- Conflict with following vehicle
- Conflict with oncoming traffic
- Conflict with a vehicle in adjacent lane
- Conflict with merging vehicle
- Conflict with a vehicle turning across subject vehicle path (same direction)
- Conflict with a vehicle turning across subject vehicle path (opposite direction)
- Conflict with a vehicle turning into subject vehicle path (same direction)
- Conflict with a vehicle turning into subject vehicle path (opposite direction)
- Conflict with a vehicle moving across subject vehicle path (through intersection)
- Conflict with a parked vehicle
- Conflict with a pedestrian
- Conflict with a pedalcyclist
- Conflict with an animal
- Conflict with an obstacle/object in roadway
- Single-vehicle conflict
- Other (specify)
- Unknown conflict

Within each conflict type there were factors that precipitated the event, that contributed to the event, and that were associated with the event. These factors are grouped into pre-event maneuvers, precipitating factors, contributing factors, associated factors, and avoidance maneuvers. The example of the relationship between these factors (for a lead vehicle, near-crash events) is shown in Figure 5.1.

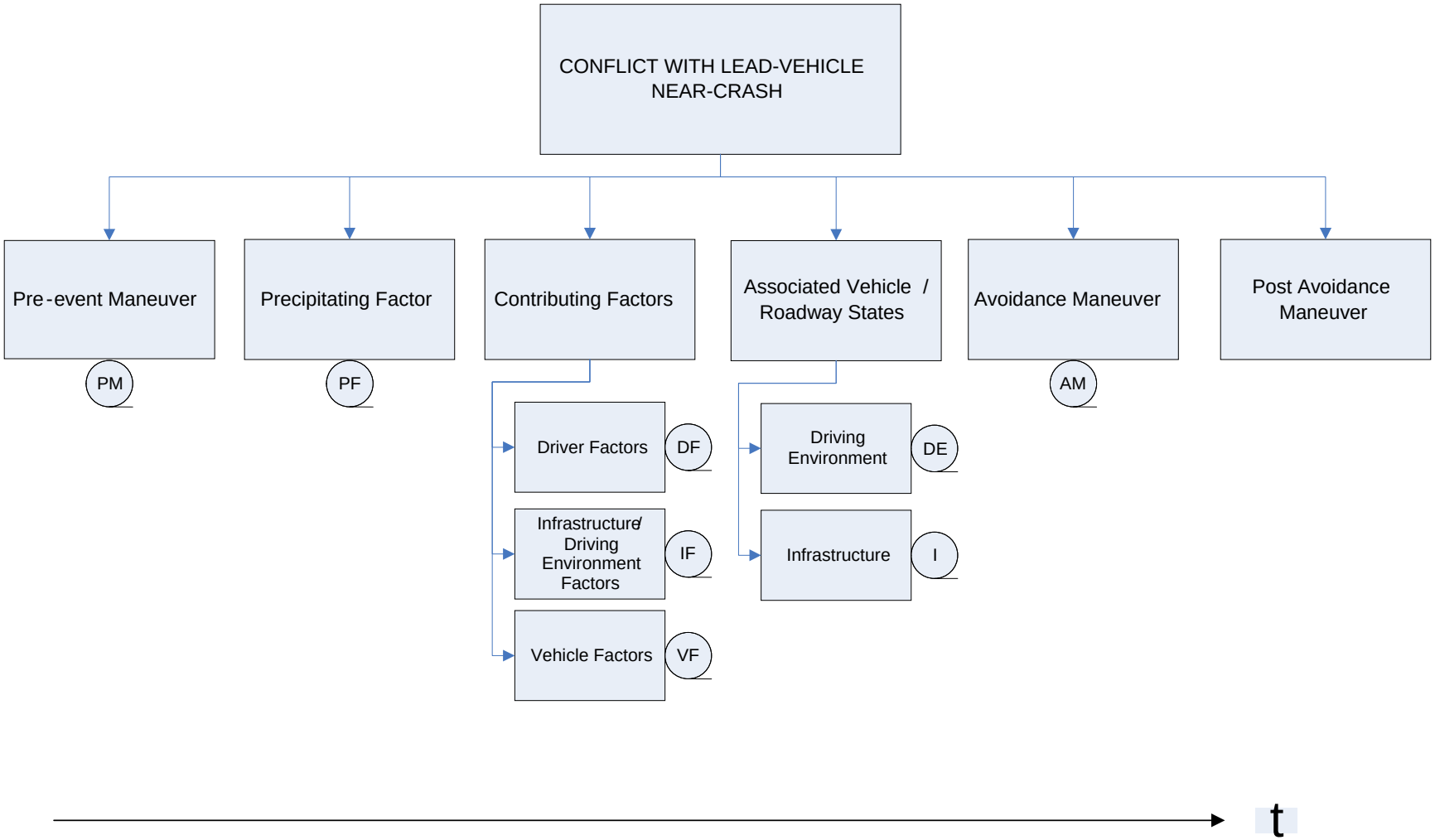


Figure 5.1. Example of the relationship between the analyzed factors for the 100-Car Study.

Recall from Chapter 2, *Method*, that a segment of time extending from 30 seconds prior and 10 seconds after the onset of the precipitating factor, was analyzed for each event, in order to catch any pre-event maneuvers, such as going straight at constant speed or changing lanes. The precipitating factor is the action that initiates the sequence of actions and circumstances that comprise the event. An example of a precipitating factor would be an animal in the roadway or a vehicle stopped for greater than 2 seconds in the traffic lane. The list of precipitating factor classifications for this study is shown in Table 5.1.

Table 5.1. Precipitating Factors Used to Classify 100-Car Study Events.

This Vehicle Loss of Control Due to:
001 = Blow out or flat tire
002 = Stalled engine
003 = Disabling vehicle failure (e.g., wheel fell off)
004 = Minor vehicle failure
005 = Poor road conditions (puddle, pothole, ice, etc.)
006 = Excessive speed
007 = Other or unknown reason
008 = Other cause of control loss
009 = Unknown cause of control loss
This Vehicle Traveling:
018a = Ahead, stopped on roadway more than 2 seconds
018b = Ahead, decelerated and stopped on roadway 2 seconds or less
021 = Ahead, traveling in same direction and decelerating
022 = Ahead, traveling in same direction with slower constant speed
010 = Over the lane line on the left side of travel lane
011 = Over the lane line on right side of travel lane
012 = Over left edge of roadway
013 = Over right edge of roadway
014 = End departure
015 = Turning left at intersection
016 = Turning right at intersection
017 = Crossing over (passing through) intersection
019 = Unknown travel direction
020a = From adjacent lane (same direction), over left lane line behind lead vehicle, rear-end crash threat
020b = From adjacent lane (same direction), over right lane line behind lead vehicle, rear-end crash threat
Other Vehicle in Lane:
050a = Ahead, stopped on roadway more than 2 seconds
050b = Ahead, decelerated and stopped on roadway 2 seconds or less
051 = Ahead, traveling in same direction with slower constant speed
052 = Ahead, traveling in same direction and decelerating
053 = Ahead, traveling in same direction and accelerating
054 = Traveling in opposite direction
055 = In crossover
056 = Backing
059 = Unknown travel direction of the other motor vehicle
Another Vehicle Encroaching into This Vehicle's Lane:
060a = From adjacent lane (same direction), over left lane line in front of this vehicle, rear-end crash threat
060b = From adjacent lane (same direction), over left lane line behind this vehicle, rear-end crash

threat
060c = From adjacent lane (same direction), over left lane line, sideswipe threat
060d = From adjacent lane (same direction), over right lane line, sideswipe threat
060e = From adjacent lane (same direction), other
061a = From adjacent lane (same direction), over right lane line in front of this vehicle, rear-end crash threat
061b = From adjacent lane (same direction), over right lane line behind this vehicle, rear-end crash threat
061c = From adjacent lane (same direction), other
062 = From opposite direction over left lane line.
063 = From opposite direction over right lane line
064 = From parallel/diagonal parking lane
065 = Entering intersection—turning in same direction
066 = Entering intersection—straight across path
067 = Entering intersection – turning into opposite direction
068 = Entering intersection—intended path unknown
070 = From driveway, alley access, etc – turning into same direction
071 = From driveway, alley access, etc – straight across path
072 = From driveway, alley access, etc – turning into opposite direction
073 = From driveway, alley access, etc – intended path unknown
074 = From entrance to limited access highway
078 = Encroaching details unknown
Pedestrian, Pedalcyclist, or other Non-Motorist:
080 = Pedestrian in roadway
081 = Pedestrian approaching roadway
082 = Pedestrian in unknown location
083 = Pedalcyclist/other non-motorist in roadway
084 = Pedalcyclist/other non-motorist approaching roadway
085 = Pedalcyclist/or other non-motorist unknown location
086 = Pedestrian/pedalcyclist/other non-motorist—unknown location
Object or Animal:
087 = Animal in roadway
088 = Animal approaching roadway
089 = Animal unknown location
090 = Object in roadway
091 = Object approaching roadway
092 = Object unknown location
099 = Unknown critical event

The *associated factors* provide a description of the driving environment and infrastructure that surrounds the event but were not judged by the trained reductionists to contribute to that event. The *infrastructure* category includes the factors that were fixed and did not change with the environment. The *infrastructure* category was further separated into the following 5 categories:

- Trafficway flow, including items such as one-way traffic and divided roadway.
- Traffic control device, including items such as traffic signal and yield sign.
- Locality, including items such as interstate and residential areas.
- Roadway alignment or road profile, including items such as straight, level, curve, and hillcrest.
- Relation to junction, including items such as intersection and entrance/exit ramp.

Driving environment consists of conditions that change on a daily or hourly basis. The traffic or driving environment is further separated into the following four categories:

- Surface condition, including wet and snowy.
- Lighting, including conditions such as streetlamps and daylight.
- Traffic density, including conditions such as stable flow, restricted speed, and restricted flow.
- Atmospheric conditions, including clear and raining.

Contributing factors were those factors that were judged by the trained data reductionists as directly influencing the presence or severity of a crash, near-crash, or incident. These contributing factors were further grouped into infrastructure/driving environment factors, driver factors, and vehicle factors. The infrastructure/driving environment factors were the same as described above as part of the associated factors, but in this case were judged as contributing to the event. For example, rain may obscure the visibility of an obstacle in the road, resulting in an event. This factor would be considered contributing and would also be included in the associated category. However, raining during a single-vehicle, run-off-road event when the driver fell asleep, would only be classified as an associated factor given that traction was not an issue.

Driver factors included willful behavior such as aggressive driving and driver impairments such as drowsiness, driver inattention, and driver proficiency errors. These driver factors provided information about any driver behaviors that most likely contributed to the severity of the event.

Vehicle factors included things such as flat tires and vehicle breakdowns. Although vehicle factors were considered in each incident, it was rarely a contributing factor, with less than 10 occurrences for all crashes, near-crashes, and incidents assessed in this study.

The factors associated with crashes, near-crashes, and incidents were extracted from the database and placed in tree diagrams. Separate tree diagrams were developed for each conflict type, event severity, and factor category. These diagrams are used to illustrate the relative frequency of each of the contributing factors for each conflict type. These diagrams include both the frequency count and the percentage to a tenth of a percentage for each factor. The percentage value is used so that comparisons for different factors between different event severities can be easily described. In the description, the percentages are rounded to the nearest percentage point. Caution should be taken when considering percentages with small frequency counts. One data point can have a large effect with a frequency count of 4, for example (i.e., 25%). Therefore the percentages should be considered along with the total frequency count when reviewing the results of this objective. A full set of the tree diagrams for all the conflicts can be found in Appendix C.

Question 1: What Are The Relative Frequencies Of Primary And Contributing Factors For Each Conflict Category?

Table 5.2 shows the relative frequency of each crash, near-crash, and incident for each conflict type. As stated earlier, there were a total of 69 crashes, 761 near-crashes, and 8,295 incidents for which data could be completely reduced. The first 8 conflict types shown in Table 5.2 accounted for all of the crashes, 87 percent of the near-crashes, and 93 percent of the incidents. Therefore, these 8 conflict types will be the focus of Question 1. Question 2 for this objective, which considers frequency of the primary and contributing factors in crashes, near-crashes, and incidents, will include all conflicts types.

The factors for each of the 8 conflict types will be described for each of the three levels of severity (i.e., crash, near-crash, and incident). The focus will be on the precipitating factor, contributing factors, and the avoidance maneuver. However, the pre-event maneuver will be discussed when it is relevant to a conflict type along with some of the associated factors.

Note that for the purpose of this objective, the factors are grouped together for each severity level of each conflict type. This gross grouping does not allow detailed deciphering of the chain of specific factors that led to a specific event. Some of the later objectives provide this detailed analysis for some categories of conflicts.

Table 5.2. Number of crashes, near-crashes, and incidents for each conflict type.

Conflict Type	Crash	Near-crash	Incident
Single vehicle	24	48	191
Lead vehicle	15	380	5783
Following vehicle	12	70	766
Object/obstacle	9	6	394
Parked vehicle	4	5	83
Animal	2	10	56
Vehicle turning across subject vehicle path in opposite direction	2	27	79
Adjacent vehicle	1	115	342
Other	0	2	13
Oncoming traffic	0	27	184
Vehicle turning across subject vehicle path in same direction	0	3	10
Vehicle turning into subject vehicle path in same direction	0	28	90
Vehicle turning into subject vehicle path in opposite direction	0	0	1
Vehicle moving across subject vehicle path through intersection	0	27	158
Merging vehicle	0	6	18
Pedestrian	0	6	108
Pedalcyclist	0	0	16
Unknown	0	1	3

Also, it is important to note that not all of these crashes were serious. For example, 75 percent of the single vehicle crashes were low g force physical contact or tire strikes. All 69 crashes were reviewed and parsed into the following four levels:

- Level I: Police-reported air bag deployment and/or injury.
- Level II: Police-reported property damage only.
- Level III: Non-police-reported property damage only.
- Level IV: Non-police-reported low-g physical contact or tire strike (greater than 10 mph).

Therefore, when reviewing this data the reader should keep in mind the severity of the crashes that are being described. The breakdown of crash severity by crash type is shown in Table 5.3. The individual Level I, II, and III crashes are described in more detail in Question 3 in this section.

Table 5.3. Crash type by crash severity category.

Conflict Type	Total	Level I	Level II	Level III	Level IV
Single vehicle	24	1	0	5	18
Lead vehicle	15	1	3	5	6
Following vehicle	12	2	2	5	3
Object/obstacle	9	0	1	3	5
Parked vehicle	4	0	0	2	2
Animal	2	0	0	0	2
Vehicle turning across subject vehicle path in opposite direction	2	1	1	0	0
Adjacent vehicle	1	0	0	1	0

Single Vehicle Conflicts

Single vehicle conflicts are conflicts that primarily involve a single vehicle departing the roadway and, in the case of a crash, colliding with an object. Single vehicle conflicts accounted for 35 percent of the crashes, 6 percent of the near-crashes, and 2 percent of the incidents. Of the 24 crashes, 22 were road departures to the left or right. The smaller percentage of near-crashes and incidents is likely due to the lack of a detected kinematic signature. As will be described in later sections, the trigger criteria for road departure events was purposely set to capture only the most severe cases that included an evasive maneuver primarily because it was difficult to distinguish planned driving maneuvers from road departure near-crashes.

Single Vehicle Crashes. The pre-event maneuver provides some additional insight into the single vehicle crash events. One third of the crashes were turning to the left or right as the pre-event maneuver (Figure 5.2). Another 17 percent were going straight while accelerating, and another 25 percent were going straight at a constant speed. One crash each was associated with the following pre-event maneuvers: changing lanes; making a U-turn; maneuvering to avoid a vehicle; decelerating in traffic lane; and entering a parking position. The tree diagram for single vehicle crashes is shown in Figure 5.2.

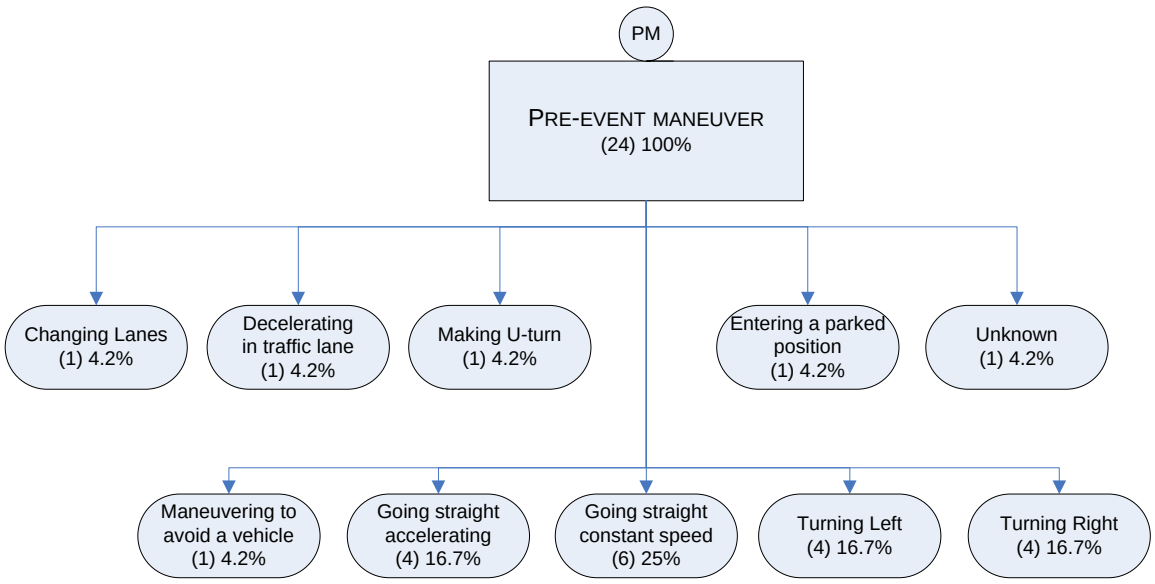


Figure 5.2. Pre-event maneuvers for single vehicle crashes

Out of the 24 single-vehicle crashes, the most common precipitating factors were the subject vehicle being over the right edge of the road (58%), over the left edge of the road (17%), loss of control due to poor road conditions (17%), and lost of control due to excessive speed (8%). Therefore, being off the edge of the road accounted for three quarters of the crashes and loss of control constituted the remaining 25 percent.

When considering driving factors, 20 percent of the single-vehicle crashes were classified as aggressive driving, 20 percent were classified as drowsiness-related, and 33 percent included driving proficiency error as a contributing factor. Inattention to the forward roadway was a factor in 46 percent of the single vehicle crashes. Of these 11 crashes, three were cell phone talking/listening, three had passengers in the vehicle, and three were attending to an object in the vehicle. The other two crashes included the driver drinking from an open container and talking/singing.

In only two crashes (8%) did drivers fail to attempt to avoid the crash. The majority (75%) steered in some manner to avoid the crash. Only 33 percent applied brakes during the avoidance maneuver. It is somewhat interesting that highest avoidance maneuver was steering to the left without braking (42%).

Infrastructure and driving environment were considered to be contributing factors in 29 percent of the single-vehicle crashes. Weather and visibility was a factor in 8 percent of the crashes. Roadway alignment was a factor in 13 percent of the crashes, and roadway delineation was a factor in the remaining 8 percent of the crashes. Glare was considered a contributing factor in two of the crashes. In one of these the glare was due to sunlight; in the other it was reflected glare. Another crash was due to a visual obstruction.

When considering other factors associated with the single vehicle crashes, 29 percent were on non-dry roads, and one-third of the crashes were at night (Figure 5.3). Two-thirds of the crashes were on straight roads, and 30 percent were on curves. One-half of the crashes were intersection-related.

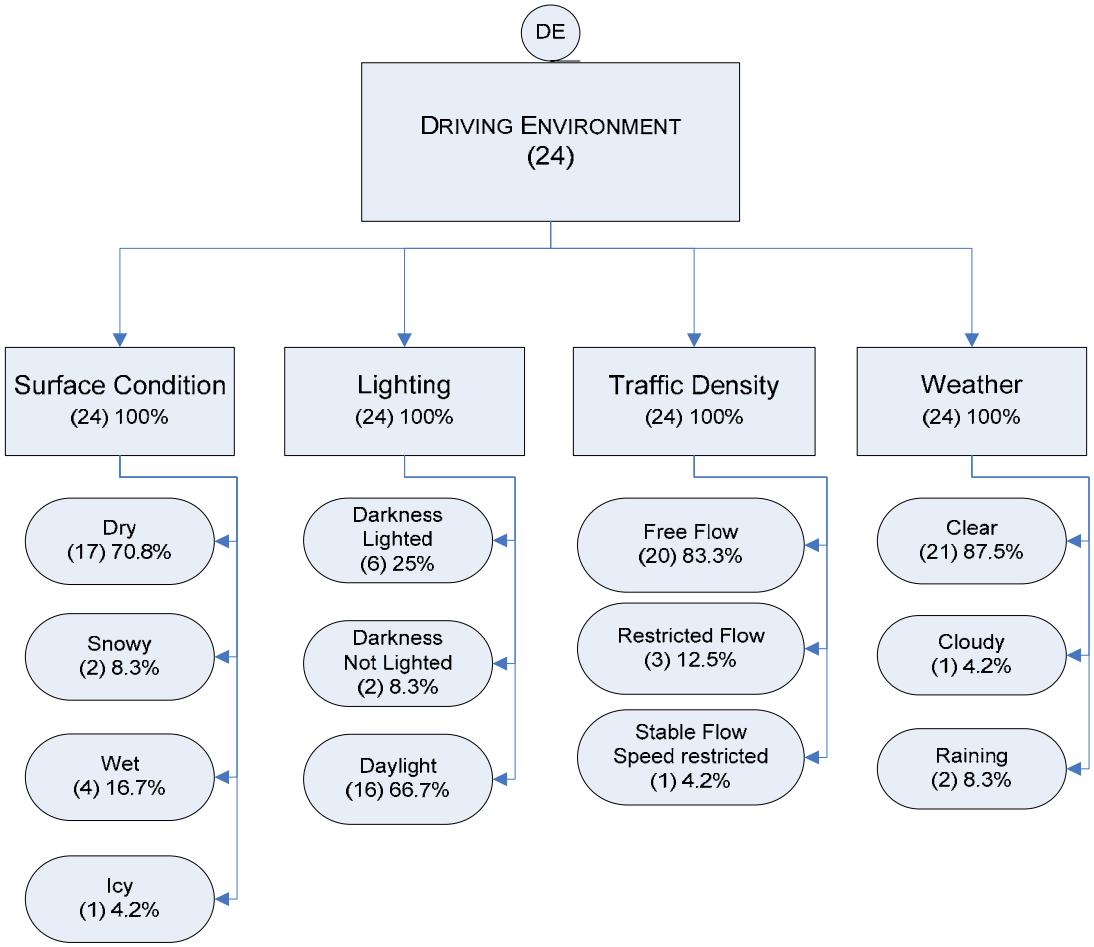


Figure 5.3. Breakdown of driving environment variables for single vehicle crashes.

Single-Vehicle Near-Crashes. There were 48 single-vehicle near-crashes identified in this analysis. As in the single-vehicle crashes, the majority of the drivers' pre-event maneuver were going straight, with 13 percent accelerating and 50 percent maintaining a constant speed. Six percent were turning left as the pre-event maneuver, and the remaining 17 percent were negotiating a curve.

Although excessive speed was a factor in 8 percent of the crashes, it was not a factor in any of the near-crashes (Figure 5.4). The most common precipitating factor was running off of the road (81%) followed by loss of control due to poor road conditions (15%).

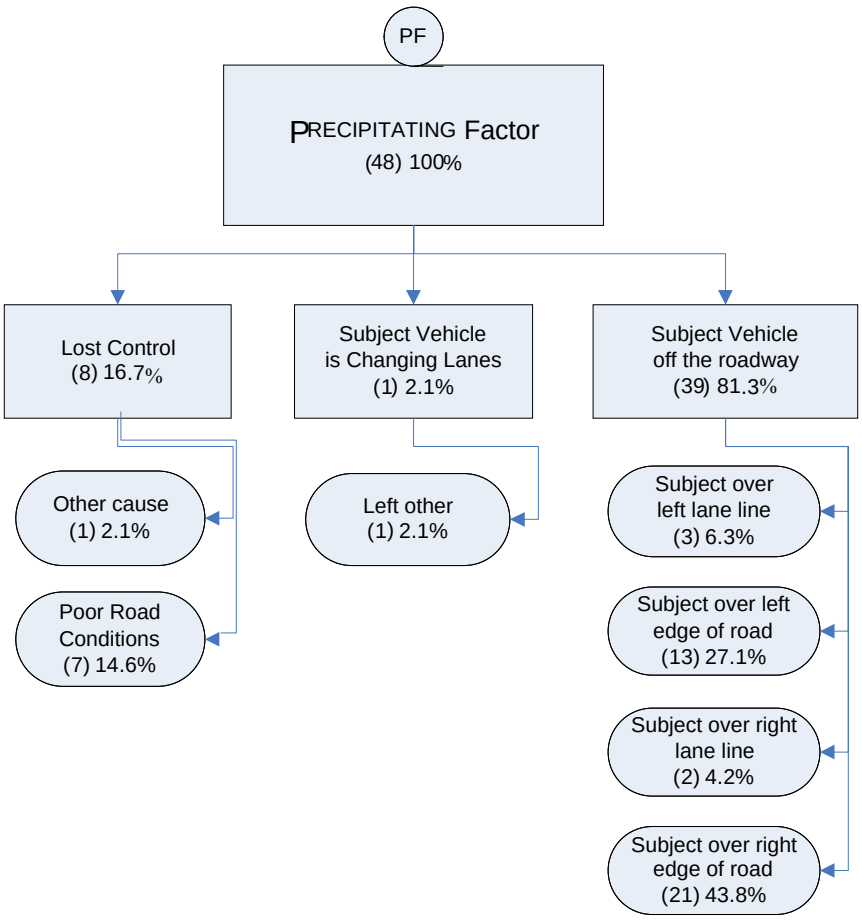


Figure 5.4. Breakdown of precipitating factors for near-crashes involving a single vehicle.

When considering driver factors, aggressive driving (4%) appeared to be less of a problem than in the crashes, whereas driver proficiency (50%) appeared more of a problem. Drowsiness (23%) was relatively the same between crashes and near-crashes.

Inattention to the forward roadway was a factor in over half of the near-crashes. Cell phone use (15%), internal, not vehicle-related distractions (10%), and vehicle-related system use (10%) accounted for the majority of the secondary task distraction (Figure 5.5). In all of these secondary tasks, it is most likely that eyes off the forward roadway contributed to the event, even for wireless device use as dialing and cell-phone other account for .

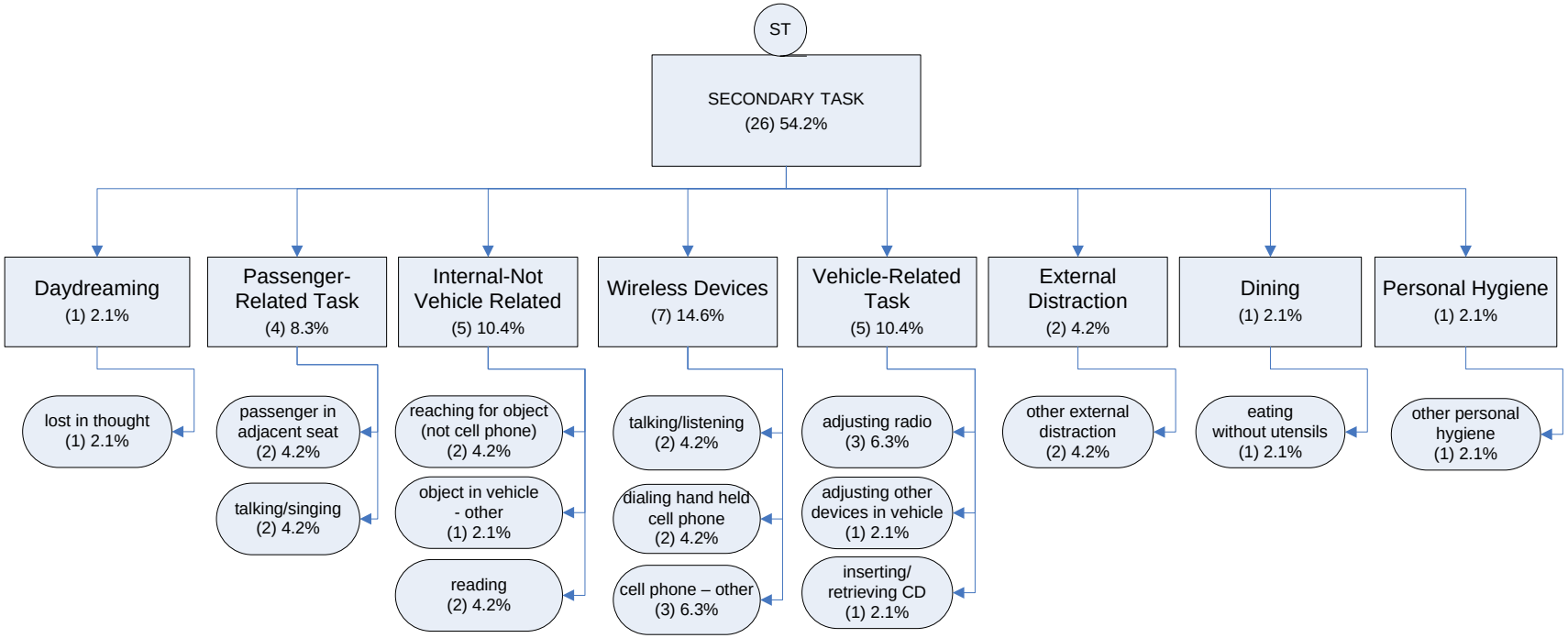


Figure 5.5. Breakdown of secondary task distractions for near-crashes involving a single vehicle.

Since the majority of these near-crashes were caused by the vehicle going off the road, steering was the most common avoidance maneuver (90%). Most drivers steered alone, with only 27 percent combining the steering with braking. No drivers braked alone during the maneuver.

The infrastructure and driving environment were considered to be a contributing factor in 23 percent of the single vehicle near-crashes. Roadway alignment (14%) was the biggest contributor in this category. Weather and visibility was a factor in 4 percent of the near-crashes, and road sight distance was a factor in one near-crash. Glare (4%) was considered a contributing factor in two of the crashes. An additional near-crash was due to a visual obstruction.

For the associated factors, 19 percent were on non-dry roads, and 54 percent were classified as during daylight (Figure 5.6). Surprisingly, 46 percent of the near-crashes occurred on curves, which is a large percentage when considering the high percentage of straight roads in the Northern Virginia/Washington, DC, area. (Figure 5.7). Higher traffic density was not associated with 88 percent of the near-crashes occurring in free-flow conditions. Intersection or intersection-related was associated with 23 percent of the near-crashes.

Single-Vehicle Incidents. There were 191 single-vehicle incidents. Similar to the crashes and near-crashes, going straight accounted for over 60 percent of the pre-event maneuvers, with 13 percent accelerating and 44 percent at a constant speed.

The most frequent precipitating factors were drivers going off the road (42%) and loss of control (41%) (Figure 5.8). The loss of control was much higher in the incidents than in the crashes and near-crashes. This loss of control was due to excessive speed (8%) and poor road conditions (16%). Although less of a factor in either the crashes or the near-crashes, turning in an intersection was a precipitating factor in 9 percent of the incidents.

Driver proficiency (63%) played a big role in the incidents, with aggressive driving (16%) and drowsiness (16%) also being contributing factors. Inattention to the forward roadway (34%) was less of a factor in the incidents than in the crashes and near-crashes. The inattention was fairly uniform across the categories, with passenger-related distraction (7%) and cell phone use (8%) being represented the most (Figure 5.9). Internal, vehicle-related, and external distractions each account for approximately 4 percent of the incidents.

Nine percent of the drivers had no avoidance maneuver, and another 9 percent braked without steering. However, as with the near-crashes, the majority steered to avoid a crash. Most drivers steered alone (50%), with 22 percent combining the steering with braking. No drivers locked up the brakes during the maneuver.

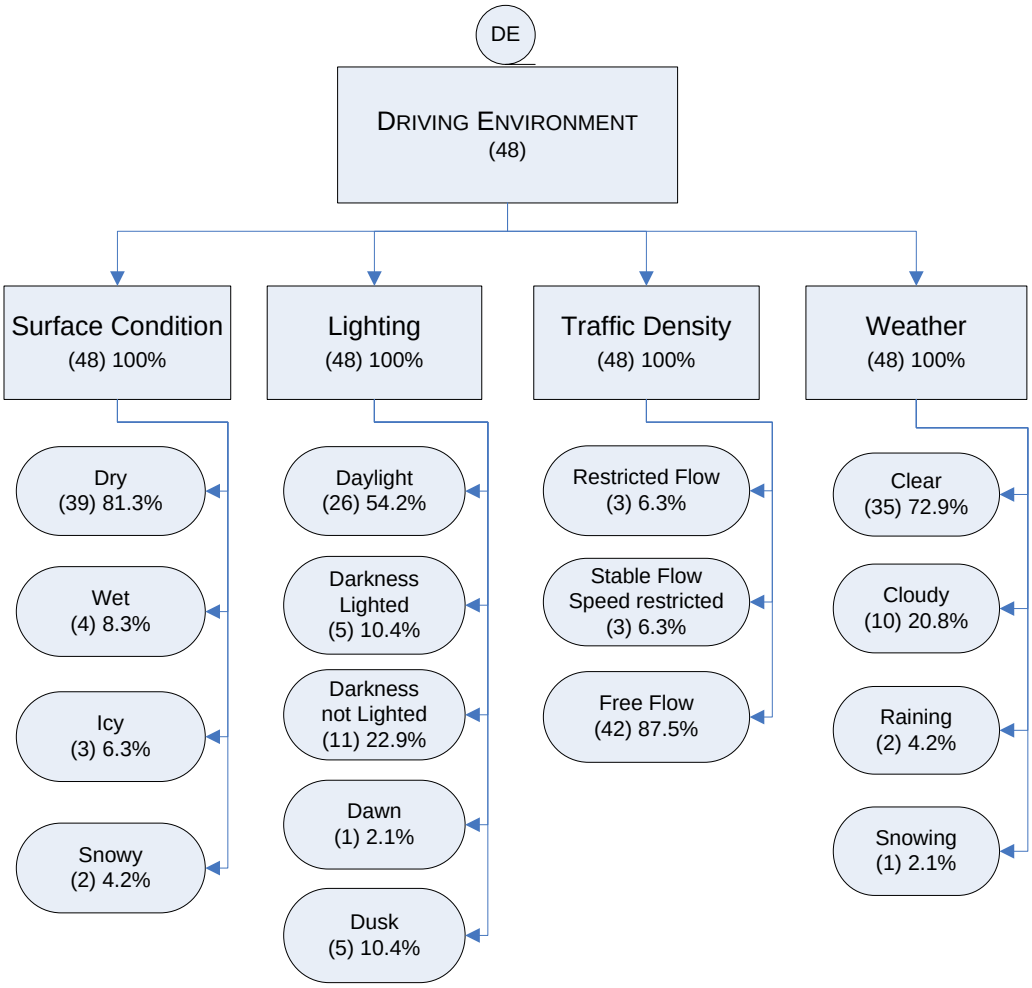


Figure 5.6. Breakdown of driving environment for incidents involving single vehicles.

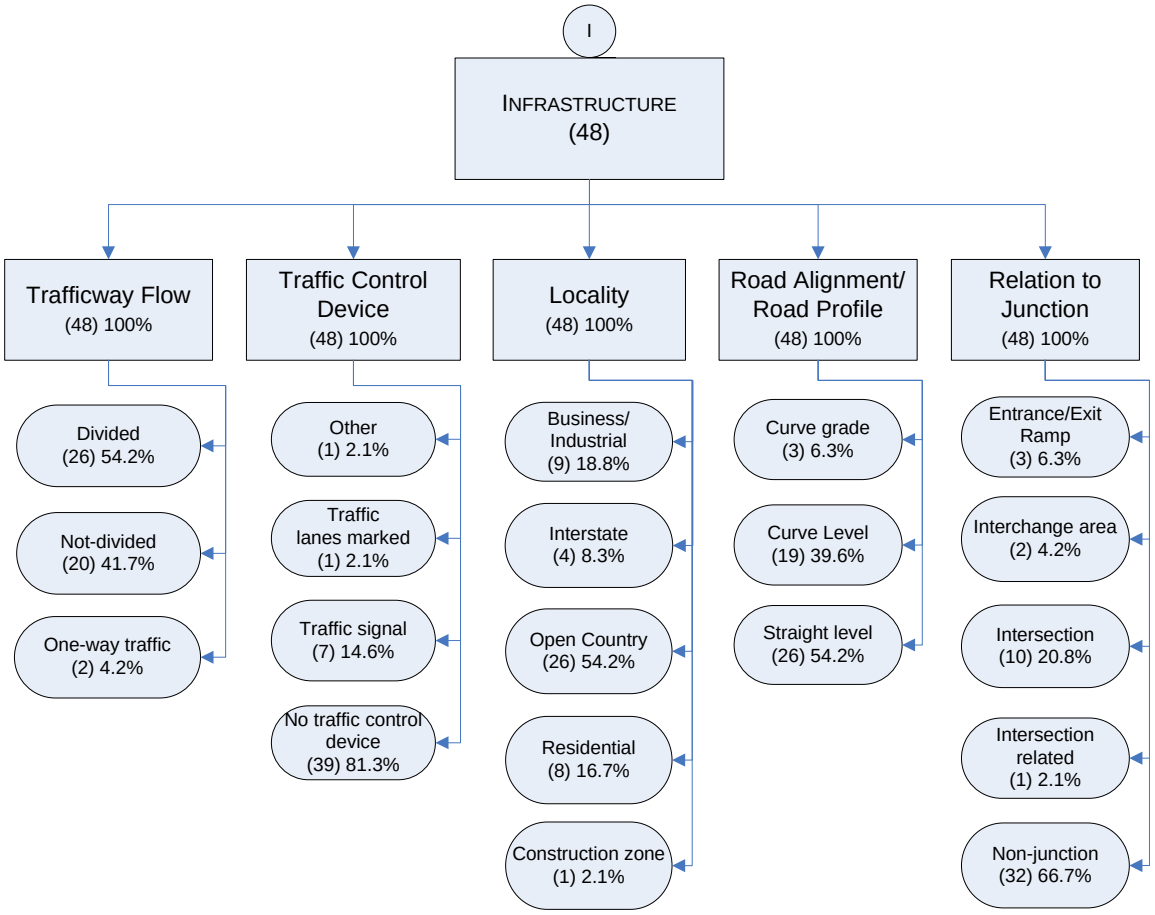


Figure 5.7. Breakdown of infrastructure variables for incidents involving single vehicles.

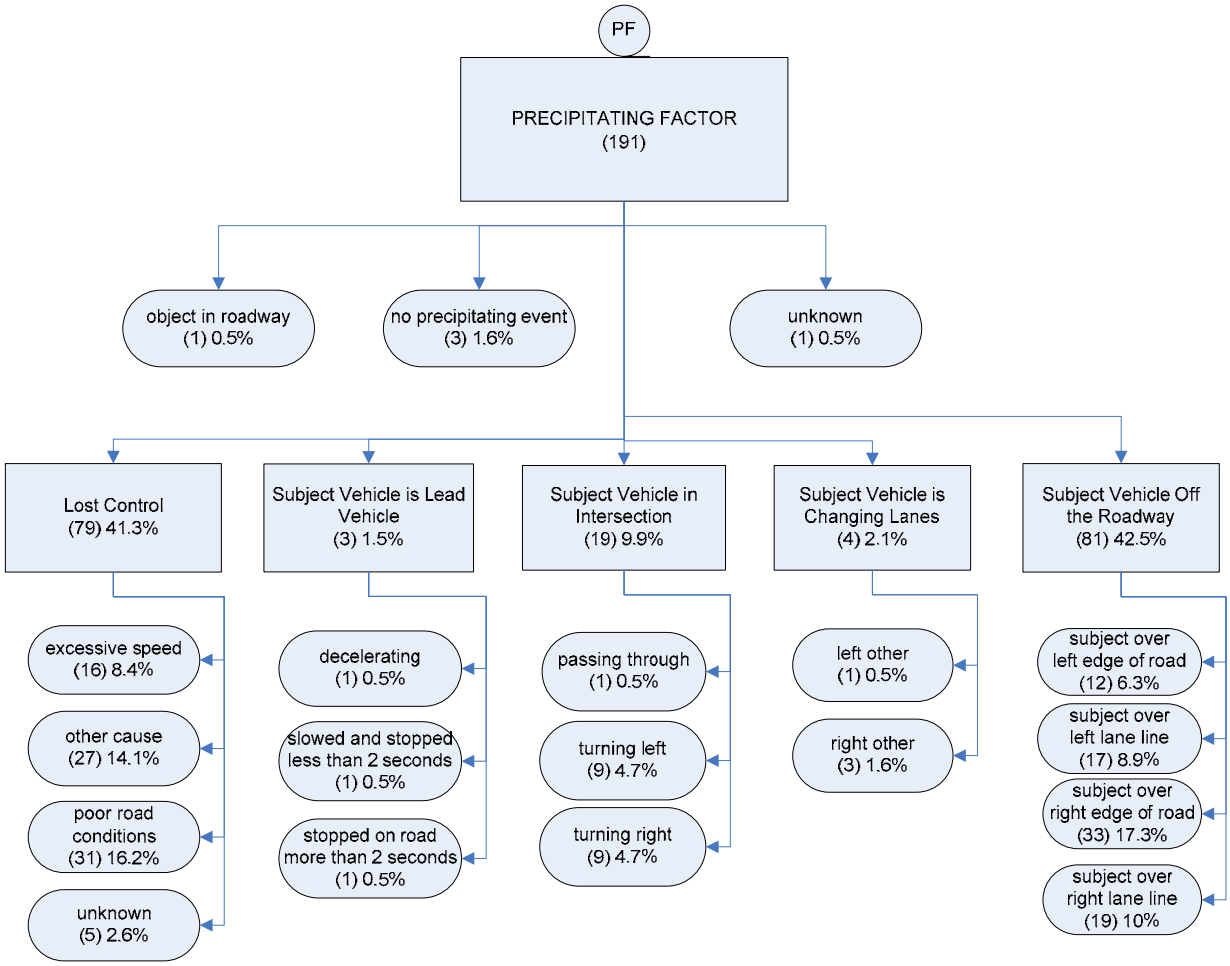


Figure 5.8. The precipitating factors for incidents involving single vehicles.

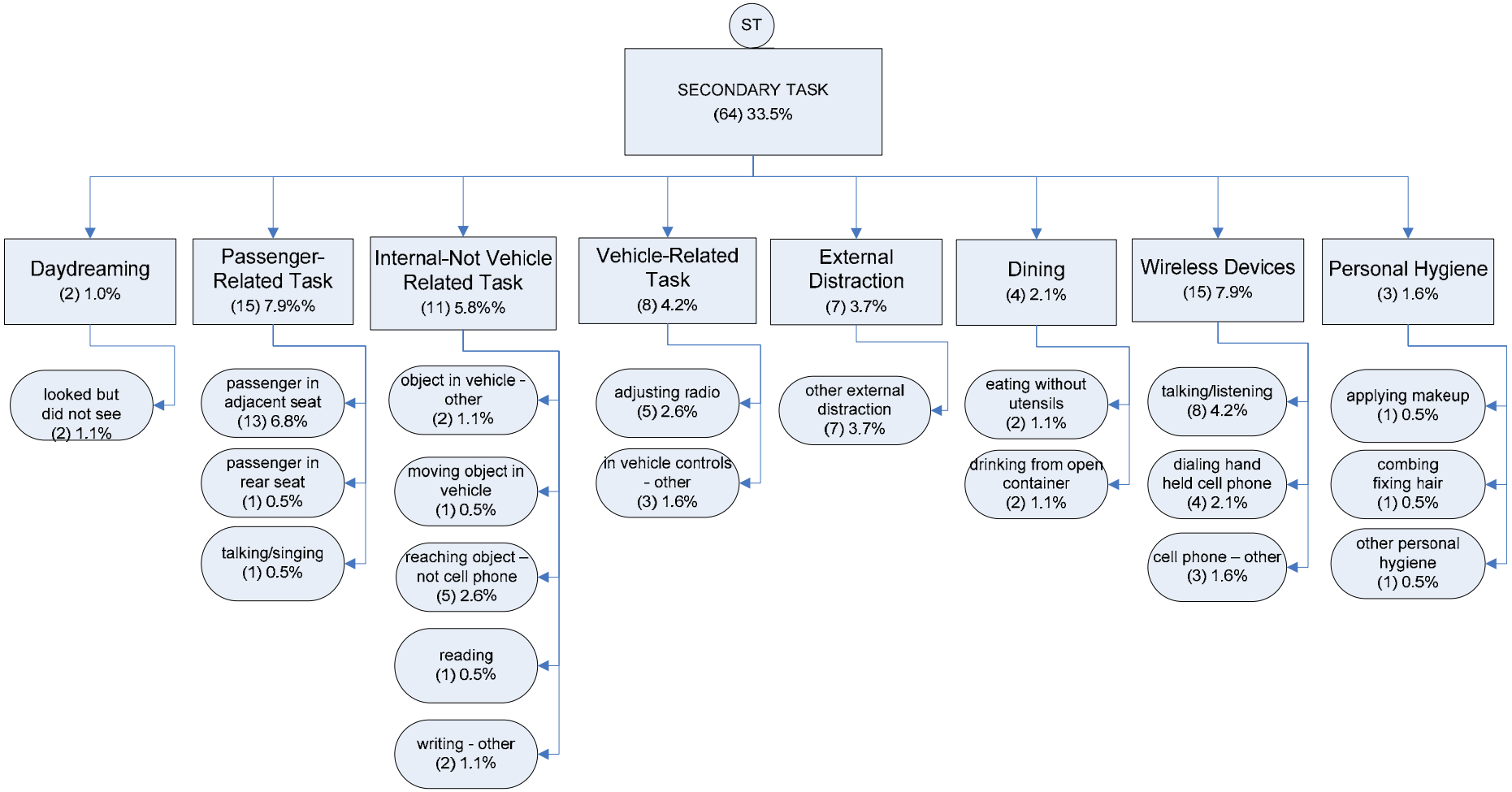


Figure 5.9. Breakdown of secondary tasks for incidents involving single vehicles.

The infrastructure and driving environment were considered to be a contributing factor in 10 percent of the single vehicle incidents. Roadway delineation (6%) was the biggest contributor in this category. Weather and visibility was a factor in 2 percent of the incidents. Roadway alignment was a factor in two incidents, and road sight distance was a factor in one incident. Glare (4%) was considered a contributing factor in 7 incidents, with 5 being due to sunlight and two being due to headlamps. An additional incident was due to visual obstruction due to a hill or curve.

For the associated factors, similar to near-crashes, 17 percent were on non-dry roads and only 57 percent were classified as during daylight (Figure 5.10). The infrastructure-associated factors indicated that 27 percent of the incidents occurred on curves and 31 percent were intersection or intersection-related (Figure 5.11). Higher traffic density was more strongly associated with incidents than near-crashes, with 27 percent being in restricted flow conditions.

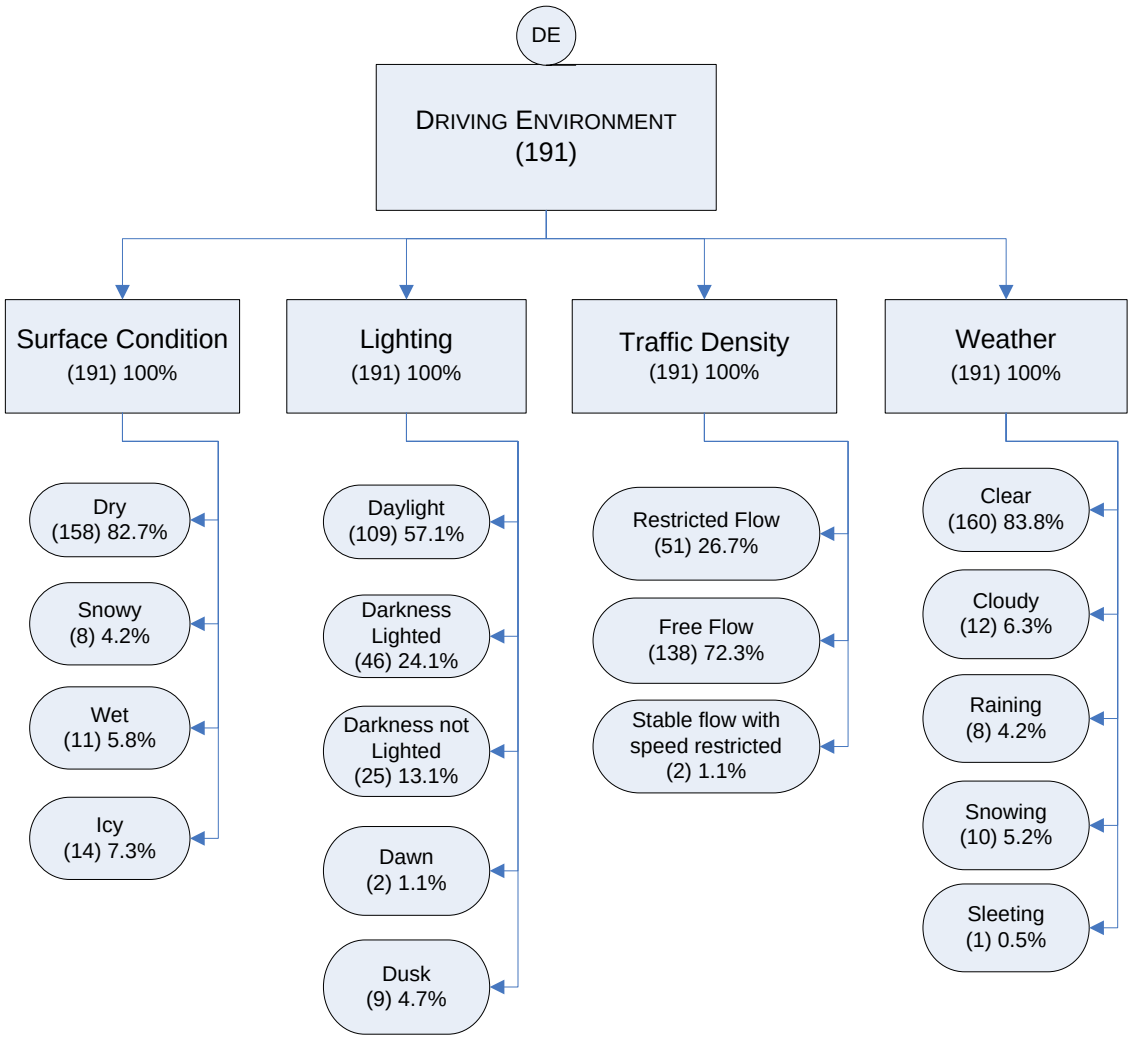


Figure 5.10. Breakdown of the driving environment variables for incidents involving single vehicles.

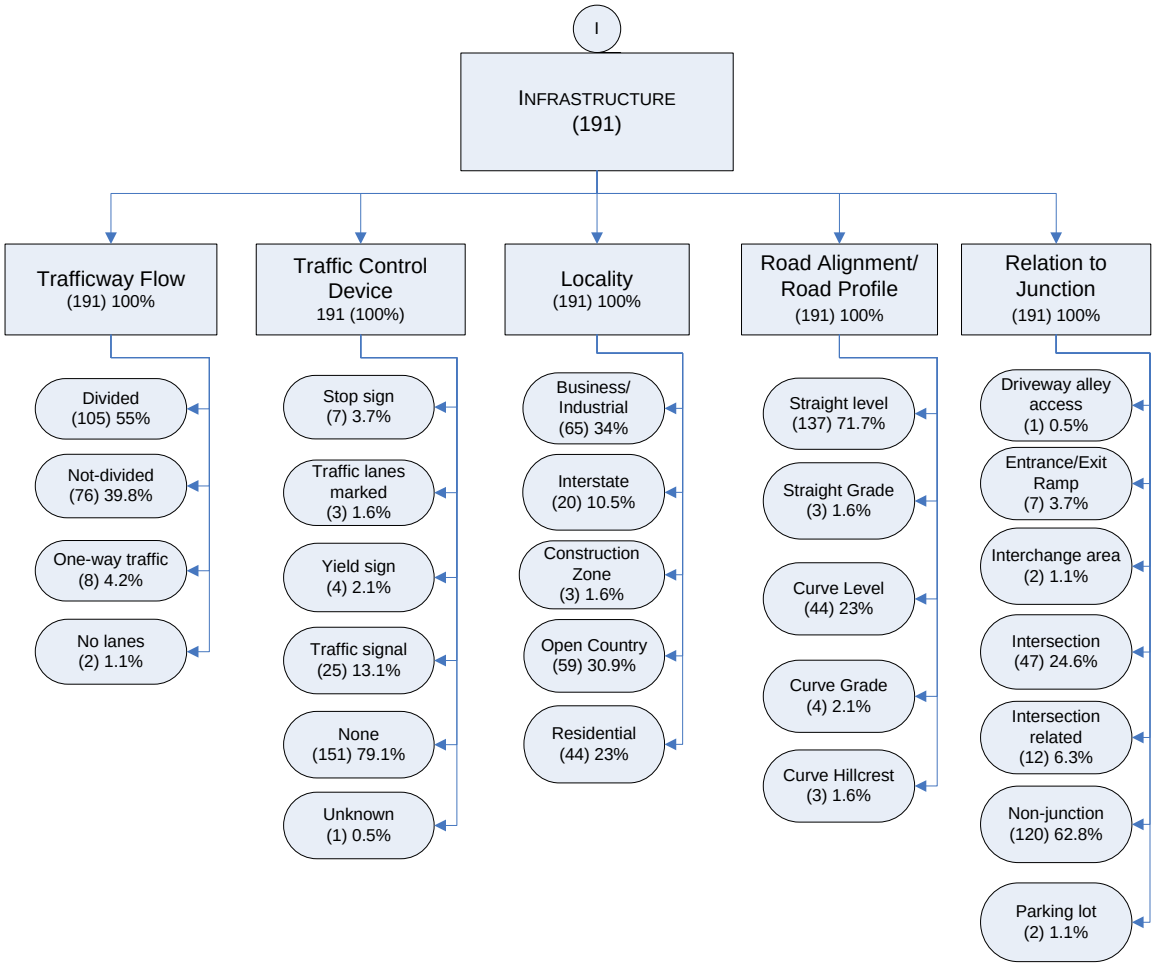


Figure 5.11. Breakdown of the infrastructure variables for incidents involving single vehicles.

Lead-Vehicle Conflicts

Lead-vehicle conflicts involve those events when an interaction occurs between the subject vehicle and the vehicle directly in front of the subject vehicle. Lead-vehicle conflicts accounted for 22 percent of the crashes, 50 percent of the near-crashes, and 70 percent of the incidents. This conflict accounted for the second largest number of crashes, but by far, accounted for the largest number of near misses and incidents. As will be discussed in later sections, the large number of near-crashes and incidents in this initial database is due in part to the presence of forward radar and establishment of trigger criteria to ensure that sufficient lead-vehicle events were categorized to address the goals of interest for this report (5 of the 10 goals were rear-end crash-related). Particularly in the case of incidents, more accurate sensors, coupled with setting the triggers more liberally (to a point) can affect the number of valid events detected. Therefore, for many of the conflict types, the incident data represent only samples of the total number present.

Lead-Vehicle Crashes. Of the 15 lead-vehicle crashes, 14 were a rear-end strike, and 1 was a road departure (i.e., lead vehicle stopped in lane and subject braked, steered off-road, and hit a telephone pole.

As the precipitating factor in 7 of the 15 lead-vehicle crashes (47%), the struck vehicle was stopped for greater than 2 seconds in the traffic lane. For another 8 of these lead-vehicle crashes (53%), the struck vehicle was stopped less than 2 seconds. This stopped greater-than- or less-than-2 seconds lead way indicates that inattention played a role in at least some of these crashes. The final crash was precipitated by a lane change.

Ninety-three percent of these crashes were categorized as having inattention to the forward road as a contributing factor (Figure 5.12). In 11 of the 15, the driver's eyes were away from the forward roadway just prior to, or during the onset of, the precipitating factors. Four of the 15 were driving-related inattention, with drivers looking out the left window (20%) or mirror (7%). In another 4 of the crashes, drivers were interacting with an object in the vehicle (27%). In an additional two crashes, drivers were dining (13%). When considering other factors, two of these crashes were classified as drowsiness or drug/alcohol-related, and two were classified as having driver proficiency error. It is interesting that no cell phone-related lead-vehicle crashes were present for this study, even though cell phone-related secondary tasks was the most commonly observed secondary task across all of the incidents and the second most common for near-crashes.

For two of the lead-vehicle crashes the driver was judged to be "daydreaming" or "lost in thought." "Lost in thought" for this study was operationally defined as the driver glancing around somewhat randomly, but not dwelling upon any particular object. These cases were not particularly common, but it was apparent in these cases that the driver was actively thinking about something other than driving.

The inattention to the forward roadway discussed earlier may explain why almost half of the drivers (47%) had no avoidance reaction. Seven of the 15 drivers (47%) did brake prior to crashing as an avoidance maneuver. Only one of the 7 locked up the brake, and only one steered while braking.

Environmental factors were not judged to be a strong contributing factor, with only one crash being due to weather and visibility. This is somewhat surprising when reviewing the associated factors, which indicated that over 40 percent of the crashes included inclement weather and wet or snowy surface conditions (Figure 5.13). Not surprisingly, traffic flow was fairly strongly associated with the lead-vehicle crashes, with only 33 percent being in free flow conditions. The infrastructure associated with the crashes was straight and level in most of the crashes (87%), with one third of the crashes being intersection-related (Figure 5.14). A single crash indicated that reflected glare was a contributing factor.

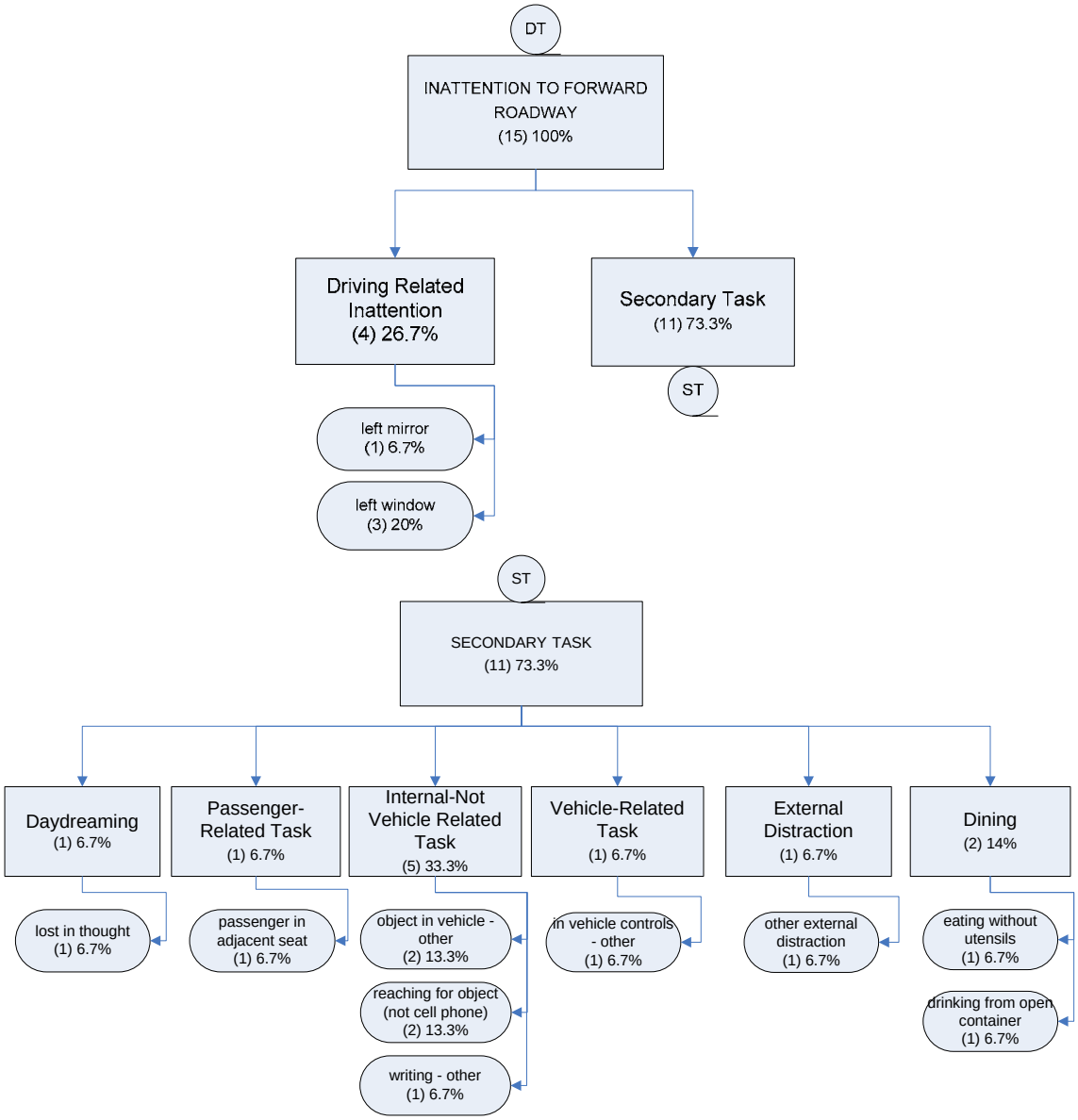


Figure 5.12. Breakdown of the secondary tasks contributing to crashes involving a lead vehicle.

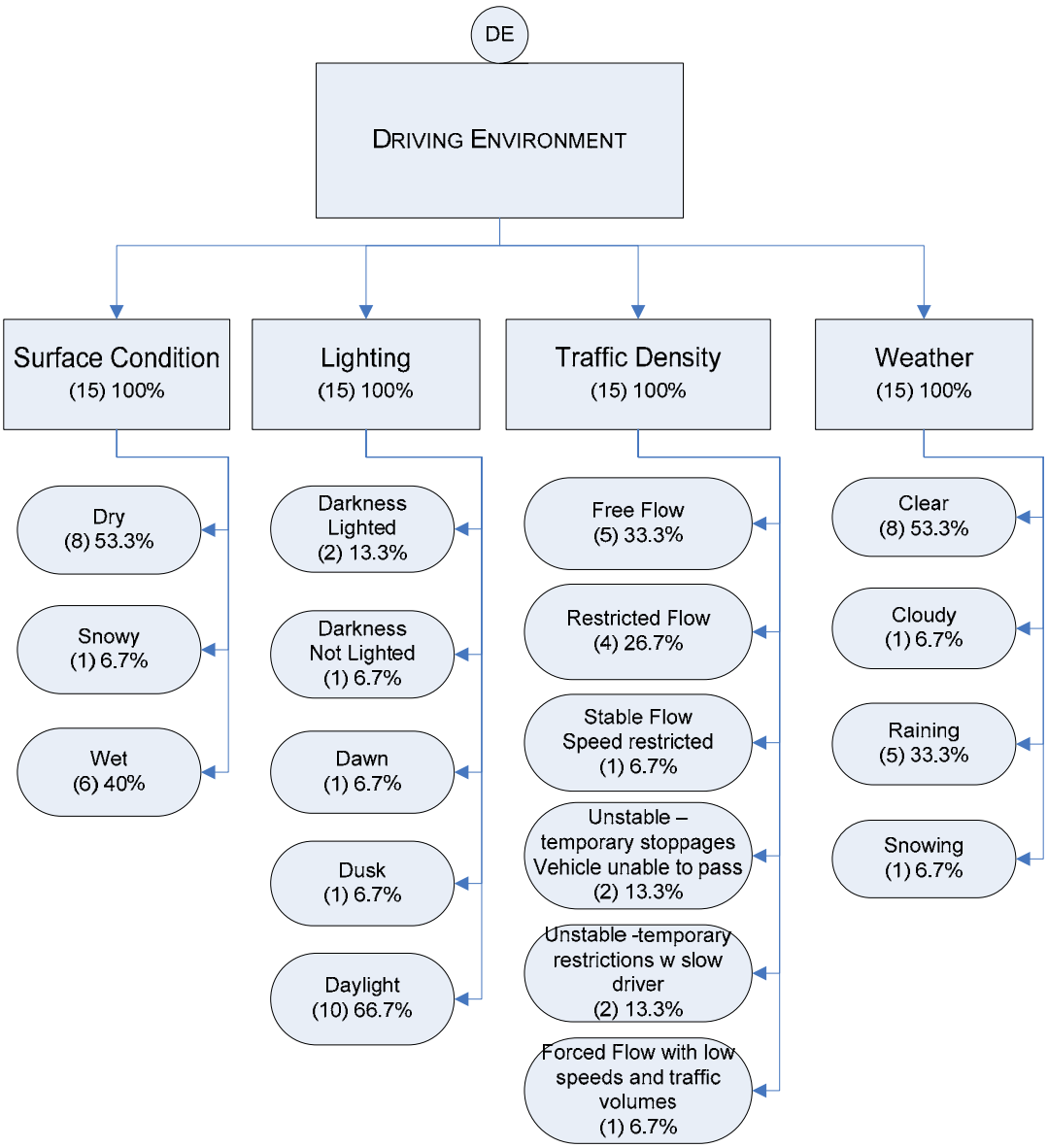


Figure 5.13. Breakdown of the driving environment variables for crashes with lead vehicles.

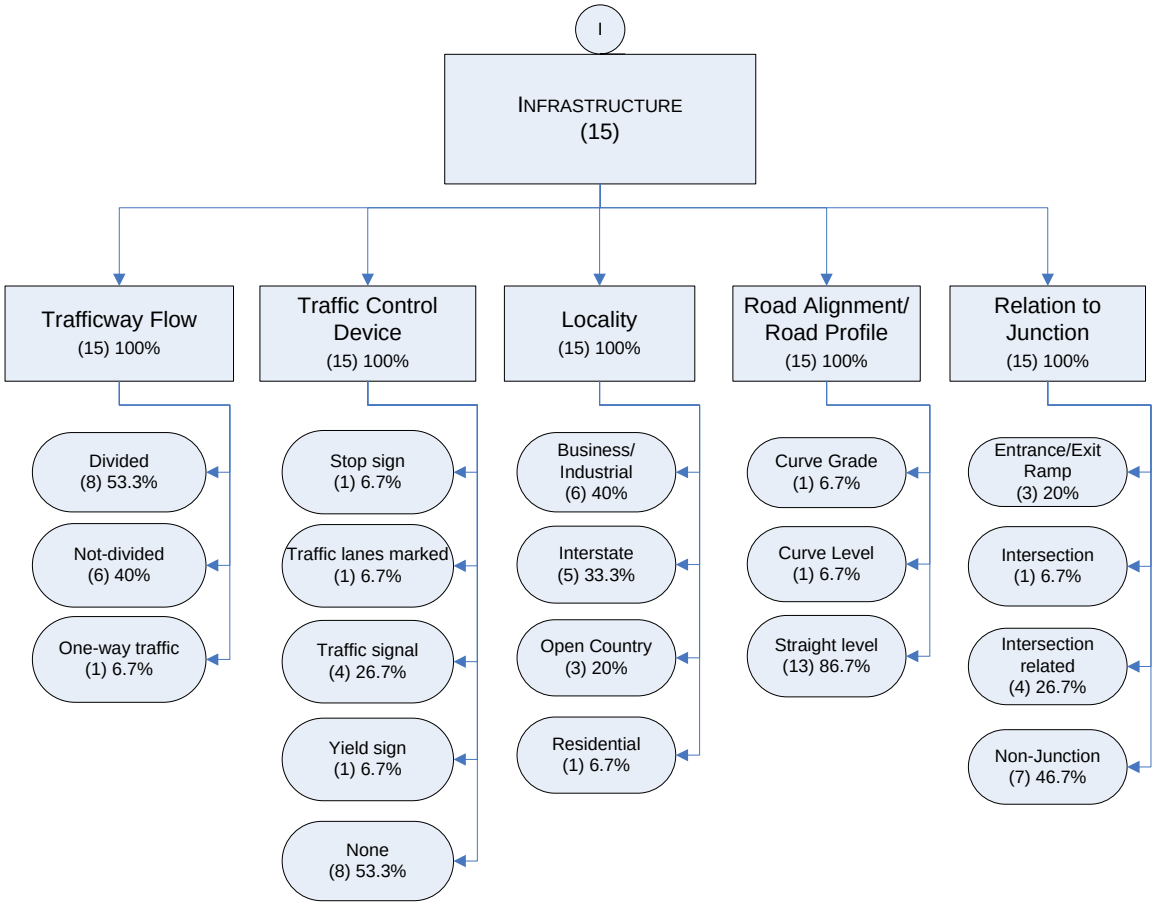


Figure 5.14. Breakdown of infrastructure variables for crashes with lead vehicles.

Lead-Vehicle Near-Crashes. There were 380 lead-vehicle near-crashes. The most common pre-event maneuvers for the lead-vehicle near-crashes were the subject vehicle decelerating in the traffic lane (21%) and going straight at a constant speed (47%). The next two most common pre-event maneuvers were the subject vehicle going straight while accelerating (16%) and the subject vehicle changing lanes (10%).

Unlike crashes, the precipitating factor associated with the lead-vehicle near-crashes was primarily lead-vehicle decelerating (42%) while lead-vehicle stopped for greater than 2 seconds and less than 2 seconds comprised 12 percent and 22 percent of the cases respectively (Figure 5.15). Finally, 18 percent of the near-crashes involved the lead vehicle changing lanes into the subject's lane of travel. These lane changes were equally representative from the left and the right.

Although still prevalent, inattention to the forward roadway was not as prevalent a factor for the near-crashes (45%) as for the crashes. Being more attentive is likely one reason some of the near-crashes did not become crashes. It may also explain why the prevalence of stopped lead-vehicle events was not as overwhelming as it was in the lead-vehicle-crash/conflict description. As described in other chapters, this study shows that drivers are more attentive when following moving vehicles at shorter headways (i.e., in "coupled" circumstances). In these cases when a driver is more attentive, a rapid deceleration by the lead vehicle in these cases would more likely result in a near-crash than a crash circumstance.

Although none of the lead-vehicle crashes had a cell phone contributing factor, cell phone use (10%) was the most frequent secondary task contributor to forward roadway inattention for near-crashes (Figure 5.16). Most of these cases were during a conversation (i.e., cell phone – talking/listening) as opposed to dialing or answering. Consistent with the above discussion, it is apparent that the cell phone conversation played a role in the event severity, but since the drivers were generally looking forward, the ultimate results, at least for the lead-vehicle conflict case, were not crashes. It is likely that this delay in reaction time contributed to near-crashes with lead vehicles, but not to the point in which the driver was unable to avoid a crash.

Driving-related inattention was a contributing factor in 13 percent of the lead-vehicle near-crashes, with drivers looking out the left window (5%), at the center mirror (3%), and out the right window (3%) being the biggest contributors. Internal distractions and not vehicle-related (6%) and passenger-related distractions (6%) were the two next most frequent contributors. The other driver factors appeared to be bigger contributing factors in near-crashes than crashes. Aggressive driving (14%), drowsiness (10%), and driving proficiency (48%) were all likely contributing factors.

Given that the operational definition of a near-crash event included an evasive maneuver, the result that all the lead-vehicle near-crashes involved an avoidance maneuver was expected. By far the most common maneuver included braking (97%). The majority of drivers braked alone (70%), but 9 percent also steered left and 18 percent also steered right. This result supports other findings (e.g., CAMP Report) that drivers braked first and then tended to steer if needed to avoid the crash.

None of the driving environment factors were identified as contributing, and only 1 percent of the infrastructure factors were identified as contributing. Three near-crashes identified road delineation as a contributing factor.

Weather was not as strongly associated with the near-crashes as with the crashes, with only 8 percent of the near-crashes including inclement weather and 12 percent including wet surface conditions (Figure 5.17). Only 21 of the near-crashes were identified as free-flow traffic, again showing the prevalence of heavy traffic as an associative factor for lead-vehicle conflicts. As in the crashes, the road was straight and level in most of the lead-vehicle near-crashes (87%). Approximately 22 percent of the lead-vehicle near-crashes were intersection-related (Figure 5.18).

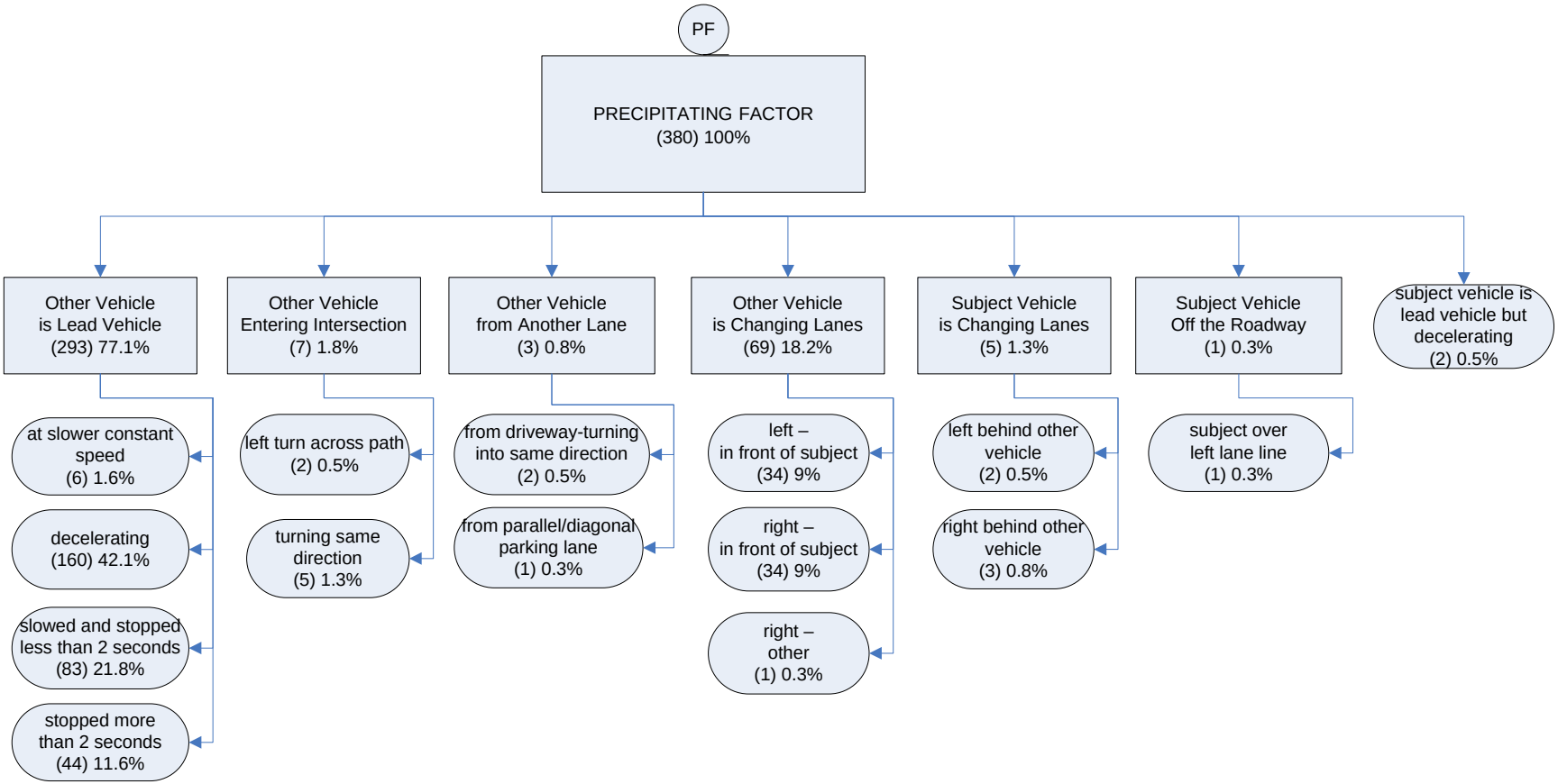


Figure 5.15. Breakdown of precipitating factor for near-crashes involving lead vehicles.

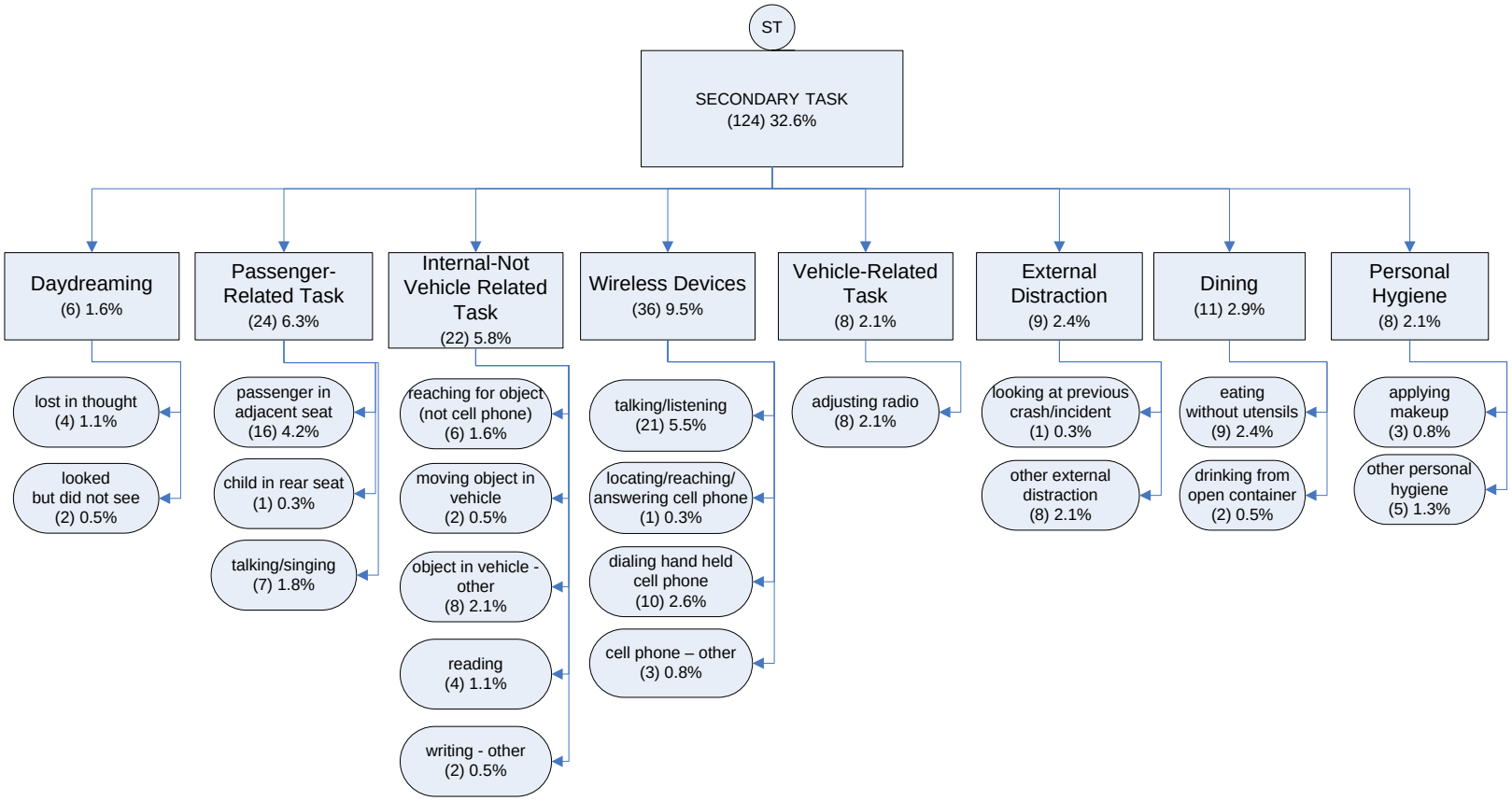


Figure 5.16. Breakdown of secondary tasks for near-crashes involving a lead vehicle.

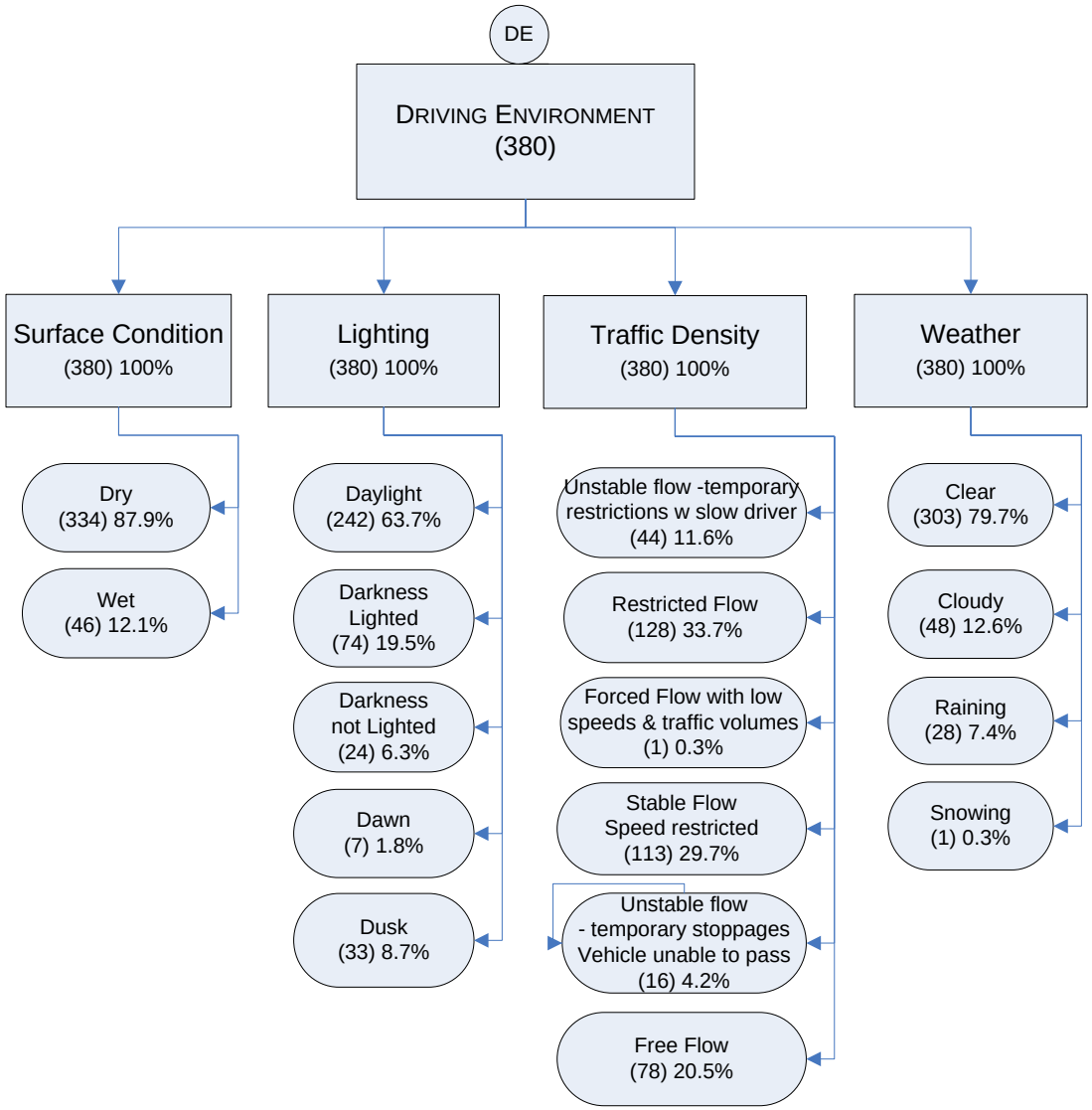


Figure 5.17. Breakdown of the driving environment variables for near-crashes involving lead vehicles.

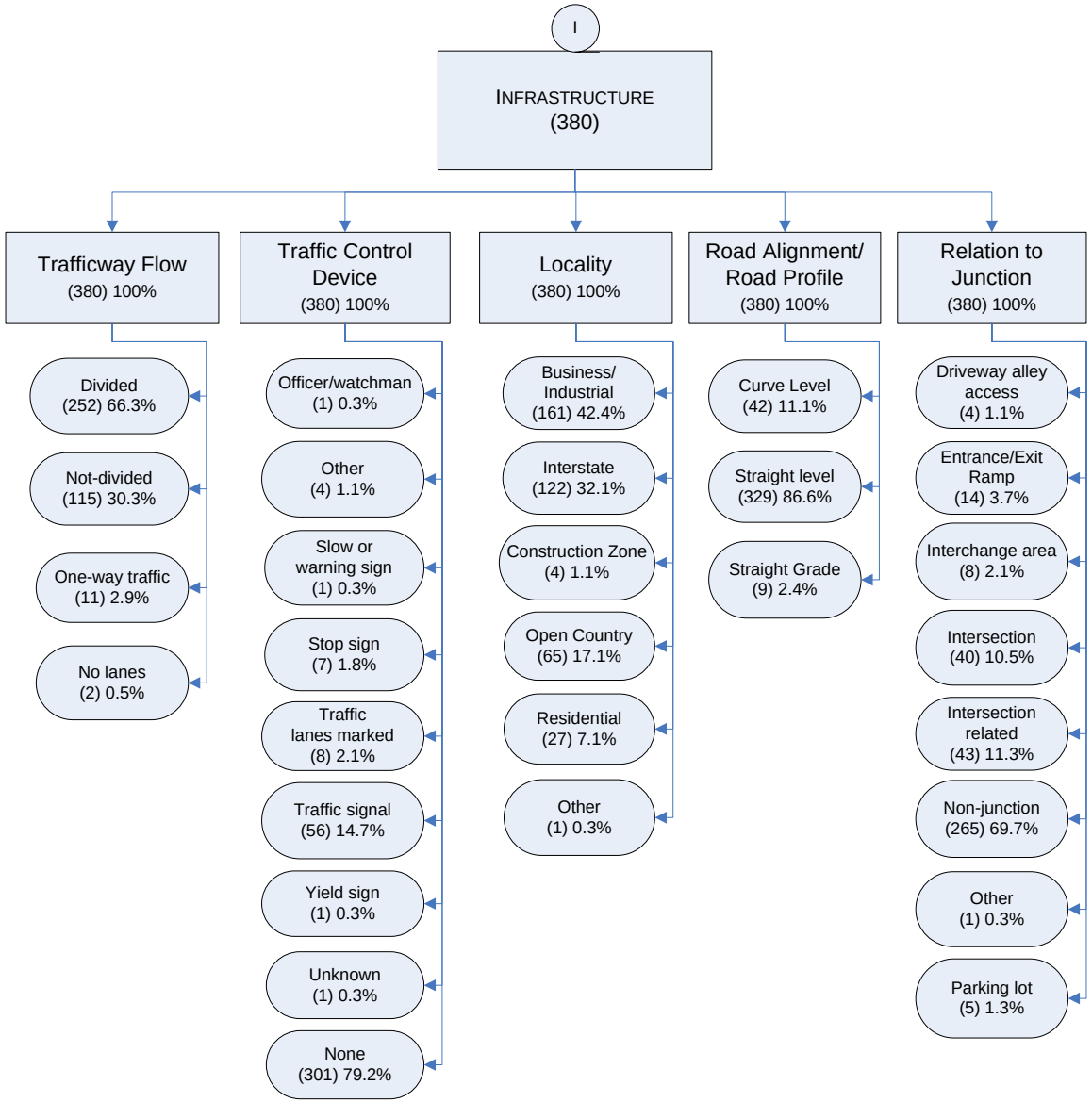


Figure 5.18. Breakdown of infrastructure variables for near-crashes involving lead vehicles.

Lead-Vehicle Incidents. As mentioned earlier, lead vehicles were involved in the largest number of incidents by far (5,783) of any of the conflicts. As described previously, this was partially attributable to the selection of trigger criteria to fulfill the objectives of these initial goals.

The most common pre-event maneuvers for the lead-vehicle incidents were the subject vehicle decelerating in the traffic lane (31%) and going straight at a constant speed (44%) (Figure 5.19). The next two most common pre-incident maneuvers were the subject vehicle going straight while accelerating (16%) and the subject vehicle changing lanes (5%).

The precipitating factor in 47 percent of the incidents was the lead vehicle decelerating. Twenty-three percent of the lead-vehicle incidents were the lead vehicle stopped for greater than 2 seconds and in another 19 percent of the incidents, the lead vehicle had been stopped less than 2 seconds (Figure 5.20). Finally, 6 percent of the incidents involved the lead vehicle changing lanes into the participant's lane of travel.

Inattention to the forward roadway was much less of a factor for the incidents than for the crashes or the near-crashes. Driving-related inattention was a contributing factor in 4 percent of the incidents, with drivers looking out the left window (2%), at the center mirror (1%), and out the right window (1%) being the biggest contributors (Figure 5.21). Although none of the lead-vehicle crashes had a cell phone contributing factor, cell phone use (8%) was the most prevalent secondary task contributor to forward roadway inattention for incidents (Figure 5.22). Talking on the cell phone accounted for 6 percent of the incidents. Passenger-related tasks and internal, not vehicle-related inattention were the next largest contributors with 5 percent and 2 percent, respectively.

As shown in Figure 5.23, driver state appeared to be a bigger contributing factor in lead-vehicle incidents than in lead-vehicle crashes. Aggressive driving (10%), drowsiness (8%), and driving proficiency (61%) were all judged as likely contributing factors.

As with near-crashes, most of the lead-vehicle incidents involved an avoidance maneuver (99%) (Figure 5.24). By far the most common maneuver included braking (95%). The majority of drivers braked alone (85%), but 4 percent also steered left, and 5 percent also steered right. These avoidance maneuvers were very similar to those seen in the near-crashes.

For the driving environment, weather was not as large of a contributing factor, with 5 percent of the incidents including inclement weather and 8 percent including wet surface conditions as associative factors (Figure 5.25). As in the crashes and near-crashes, the road was straight and level in most of the lead-vehicle incidents (92%). Approximately 30 percent of the lead-vehicle incidents were intersection-related.

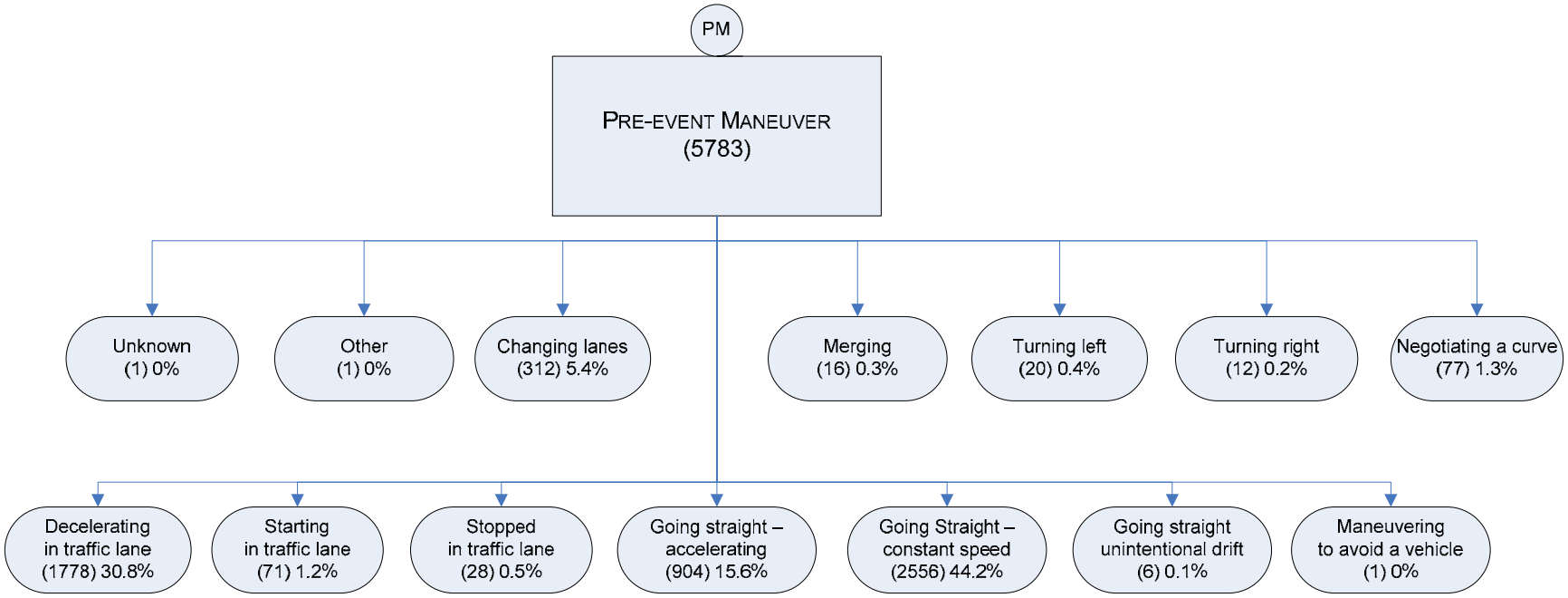


Figure 5.19. Breakdown of all pre-event maneuvers that occurred prior to incidents involving lead vehicles.

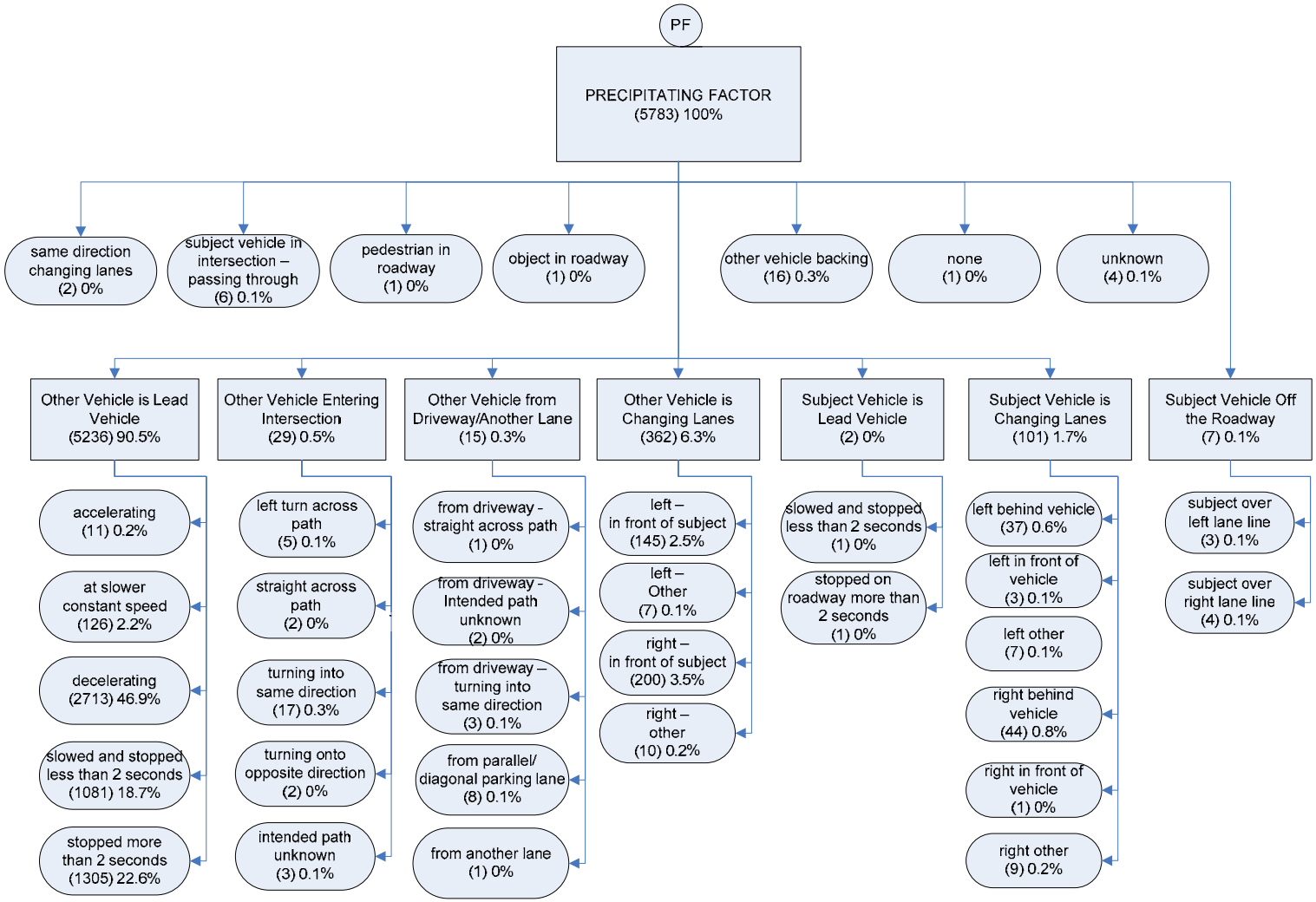


Figure 5.20. Breakdown of precipitating factors for incidents involving lead vehicles.

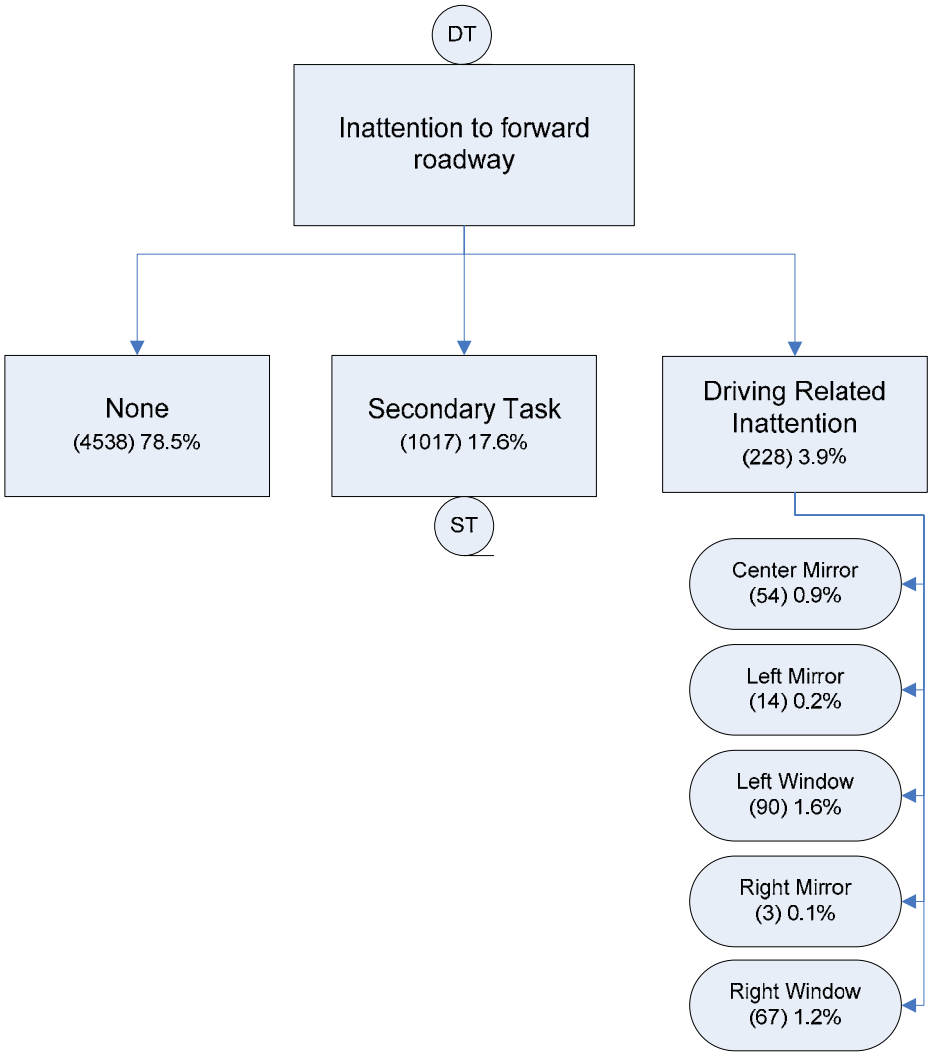


Figure 5.21. Breakdown of inattention categories for incidents involving a lead vehicle.

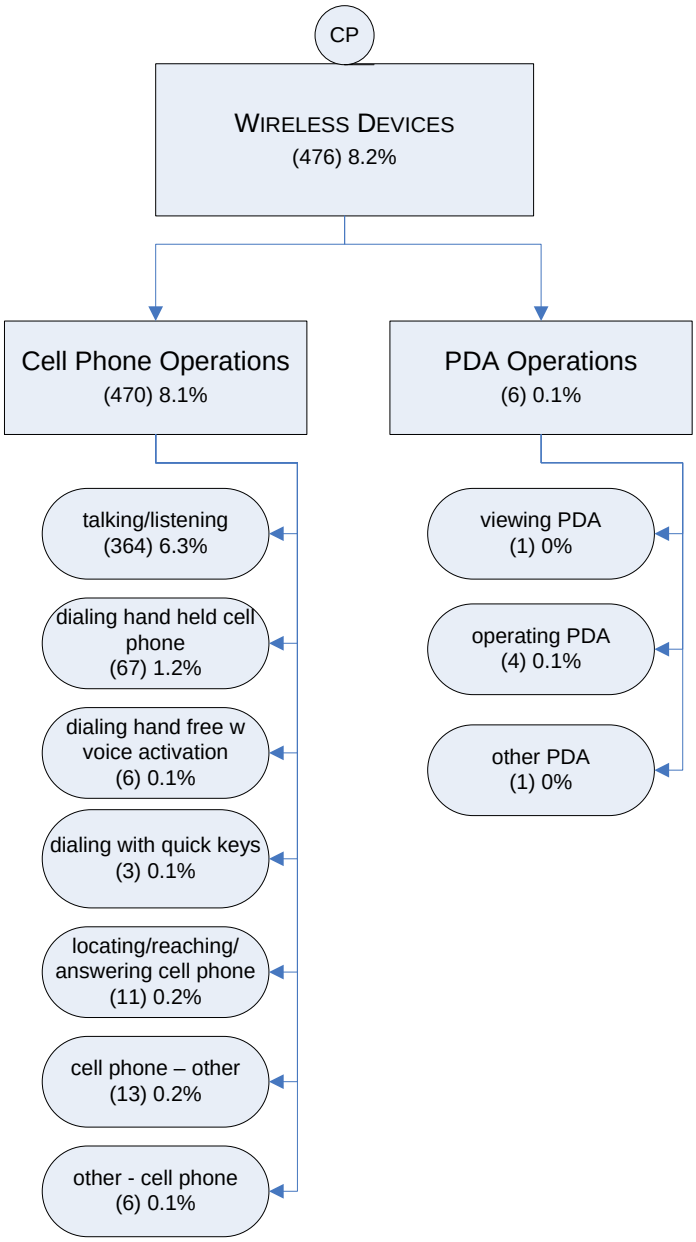


Figure 5.22. Breakdown of wireless device operations that contributed to incidents involving lead vehicles.

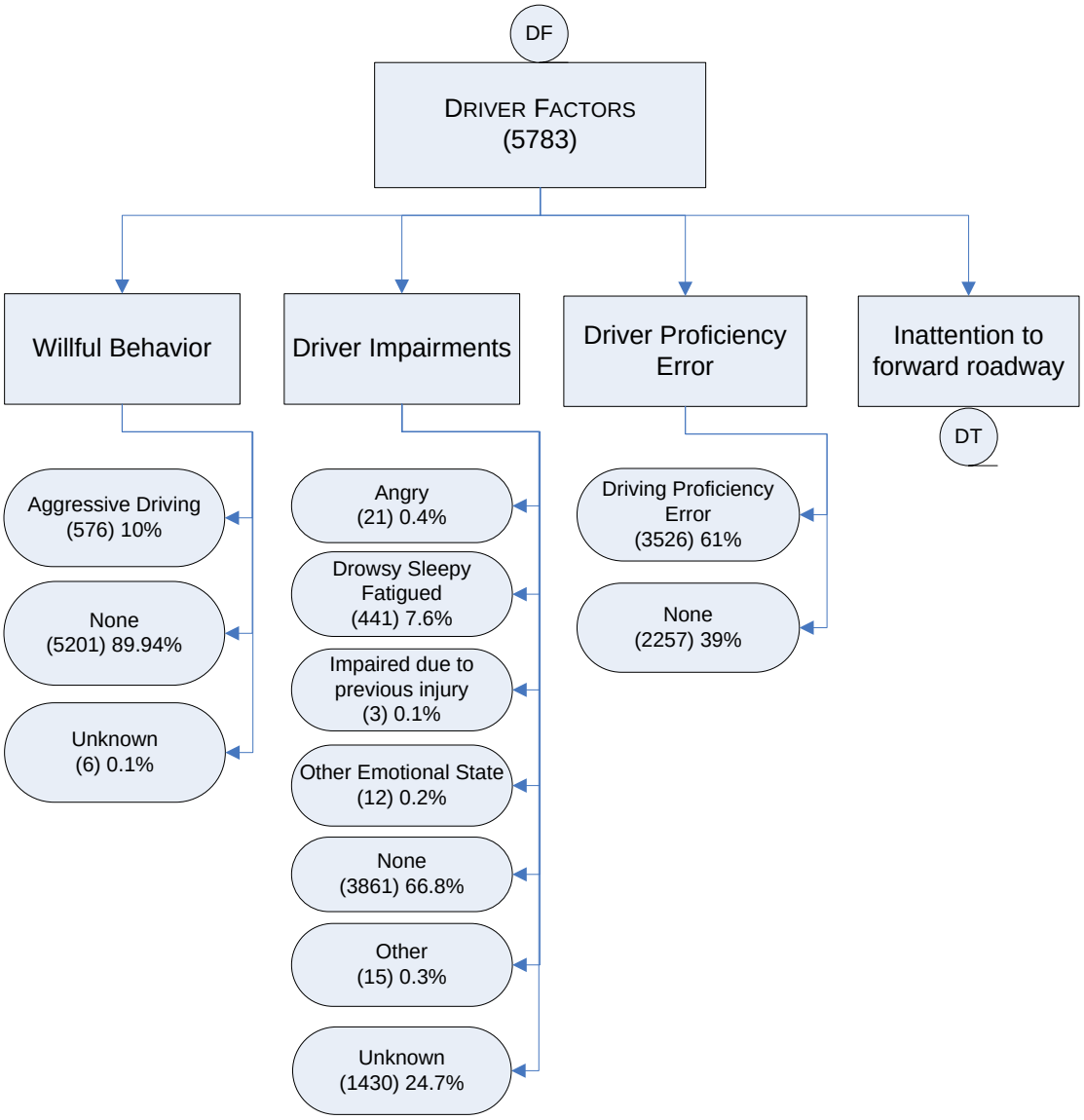


Figure 5.23. Breakdown of driver factors that contribute to incidents involving lead vehicles.

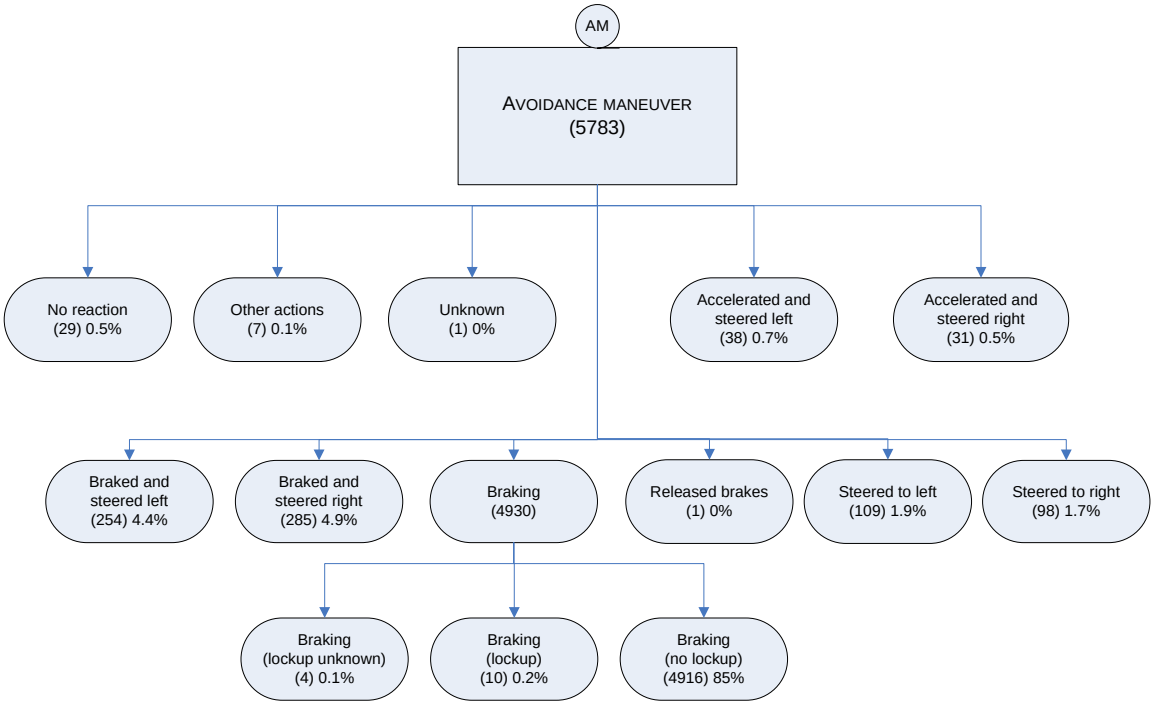


Figure 5.24. Breakdown of avoidance maneuvers that occurred during incidents involving lead vehicles.

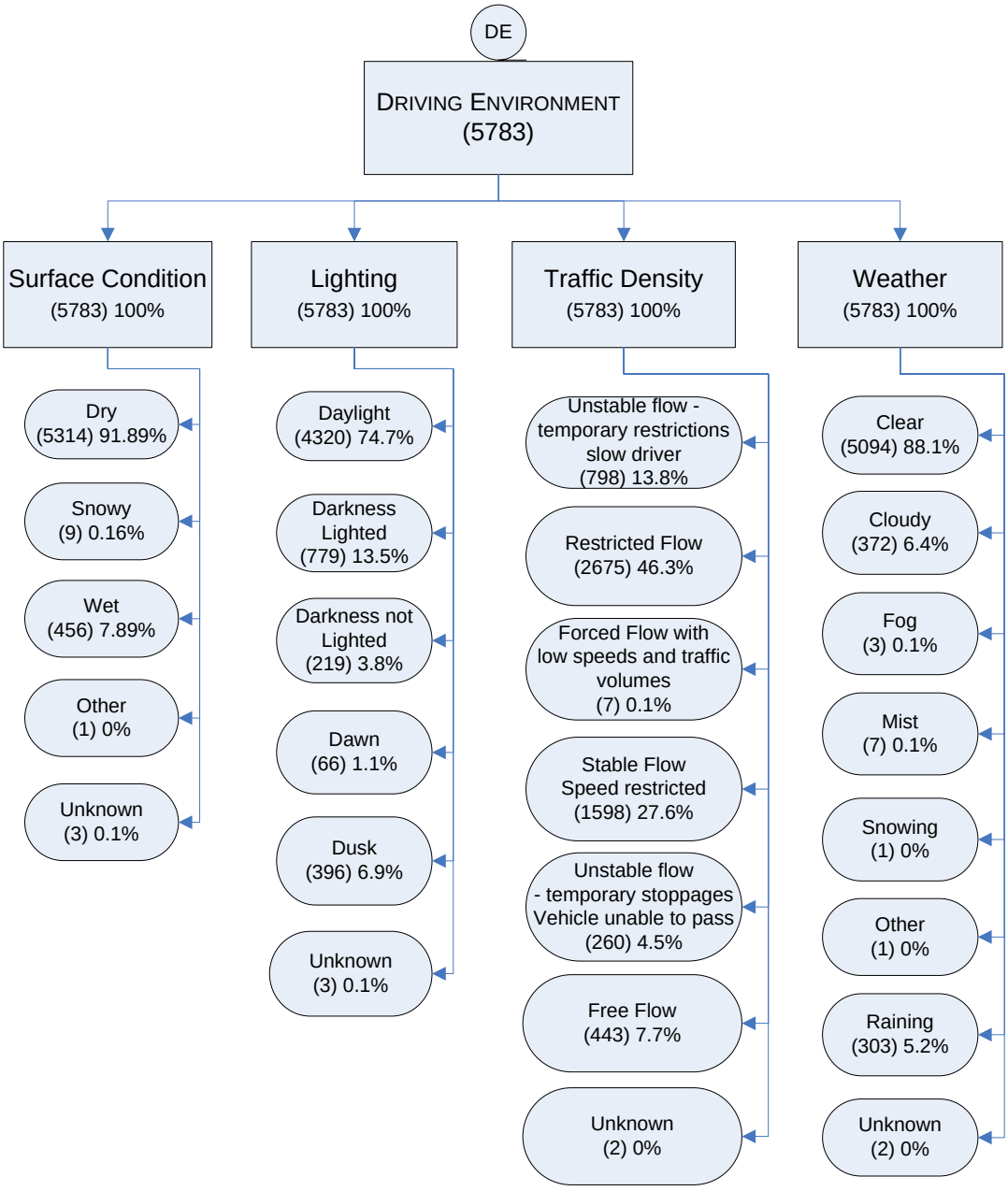


Figure 5.25. Breakdown of the driving environment variables that contribute to incidents involving lead vehicles.

Following-Vehicle Conflicts
Following vehicle conflicts are conflicts that occur with the subject vehicle and the vehicle directly *behind* the subject vehicle. Following vehicle conflicts accounted for 18 percent of the crashes, 9 percent of the near-crashes, and 9 percent of the incidents. This conflict had the third highest number of crashes with 12. All 12 crashes involved the rear end of the subject vehicle being struck. While the frequency of lead and following-vehicle conflicts should theoretically be similar, data reductionists experienced some difficulties when validating these events. The rear camera was located on the rear dash of the vehicle and was turned at an angle to capture a greater portion of the passenger-side space surrounding the vehicle. Placing the camera this far away from the driver's view-point made it difficult to gauge the severity of the event (i.e., how close the following vehicle came to the subject's vehicle). Turning the camera at an angle also presented difficulties in determining the proximity of the following vehicle. Therefore, the trigger criteria were set much more conservatively to ensure that all of the incidents and near-crashes in this category were valid.

Following-Vehicle Crashes. For the following-vehicle crashes, the events were examined in two ways – whether the following vehicle was at fault or whether our driver contributed to the event.

As the precipitating factor in 5 of the 12 following-vehicle crashes (42%), the SV was stopped for greater than 2 seconds in the traffic lane. In an additional 4 of the crashes, the SV was decelerating (33%). The remaining three crashes were precipitated by the subject vehicle stopped less than 2 seconds. With this information, the stimulus response time (SRT) of the following vehicle was calculated for 5 following-vehicle level I, II, and III crashes based on data from the rear radar (Table 5.4).

Table 5.4. Stimulus response time and crash times for following-vehicle level I, II, III crashes.

Epoch #	SRT	Crash Time	Comments
0040403061354014200	5.1 s	6.7 s	Women may have hit husband intentionally (one of two consecutive rear-end crashes with her husband)
0120404081309002412	N/R	2.3 s	FV appears to accelerate into accident. Perhaps due to inattention as the SV was not expected to stop during a left-turn yield maneuver
0180310281755018936	2.1 s	4.9 s	FV appears to start decelerating and then accelerates into crash, perhaps thinking that the heavy traffic was surging forward
0520305062218000000	N/R	1.1 s	SV was always on the brake, so point at which SV stopped was used instead. No apparent FV reaction
0820310311144005225	1.6 s	4.6 s	FV begins to decelerate quickly, but insufficiently to avoid the accident
* N/R = No crash avoidance reaction indicated by radar pattern			

SRT was defined as the time from when the SV driver contacted the brake pedal until the following vehicle (FV) first began to decelerate. The crash time was also calculated. Crash time is operationally defined as the time from when the participant first contacted the brake until the impact began. The average SRT and crash times as well as the average crash time for

participants that did not react were also calculated (Table 5.5). The no-reaction crash time was calculated because it indicates that the driver of the following vehicle was unable to react in the time prior to impact.

Table 5.5. Average stimulus response time and crash times for following-vehicle level I, II, III crashes.

N=3, Average SRT=	2.9 s
N=3, Average Crash Time w/SRT =	5.4 s
N=2, Average Crash Time w/o SRT =	1.7 s

The short average crash time without an SRT indicates that the crash happened so quickly that the driver basically could not respond. This is highlighted by the fact that the average SRT was 2.9 seconds, which was more than a second longer than the crash time without an SRT. In addition, there were three crashes for which problems associated with the rear radar precluded the response calculations. The omitted data points and the reason for omission are shown in Table 5.6.

Table 5.6. Omitted epics in SRT and crash time calculations for following-vehicle level I, II, III crashes.

Epoch #	Reason not used
0030304091846007000	Insufficient radar track. FV not picked up until only 20 feet away
0040403061354013156	Radar tracking other vehicle, FV not tracked
0630311251458003137	Rear radar malfunction, no data

The following-vehicle crashes for which the SV driver contributed was also examined. Driving proficiency was identified as a contributing factor in 33 percent of the crashes. Aggressive driving (8%) and drowsiness (8%) were contributing factors in one crash each.

Seven of the 12 drivers (58%) in the following-vehicle crashes had no reaction. In most of these cases the subject vehicle was likely stopped. Of the remaining 5 crashes one braked and steered left, and the other 4 braked alone.

None of the driving environment factors were identified as contributing, and only one crash infrastructure factor (i.e., roadway delineation) was identified as contributing.

Weather was not a large associated factor, with no inclement weather and only two wet surface associated conditions (Figure 5.26). Only 4 of the 12 crashes were in free flow conditions. As shown in Figure 5.26, roadway alignment may have played a role, with 42 percent of the crashes being on curves. Two-thirds of the crashes were intersection-related.

Following-Vehicle Near-Crashes. Not surprisingly, there were more varied precipitating factors for the 70 following-vehicle near-crashes than for the following-vehicle crashes (Figure 5.27). A third of the near-crashes in this conflict were due to a subject vehicle decelerating. In 23 percent of the near-crashes the subject vehicle was stopped for less than 2 second. Although no following-vehicle crashes were associated with the subject vehicle changing lanes, 24 percent of

the near-crashes included the subject vehicle into the traffic lane of the conflict. The maneuver was to the left in 14 percent of the near-crashes and to the right in 10 percent of the near-crashes.

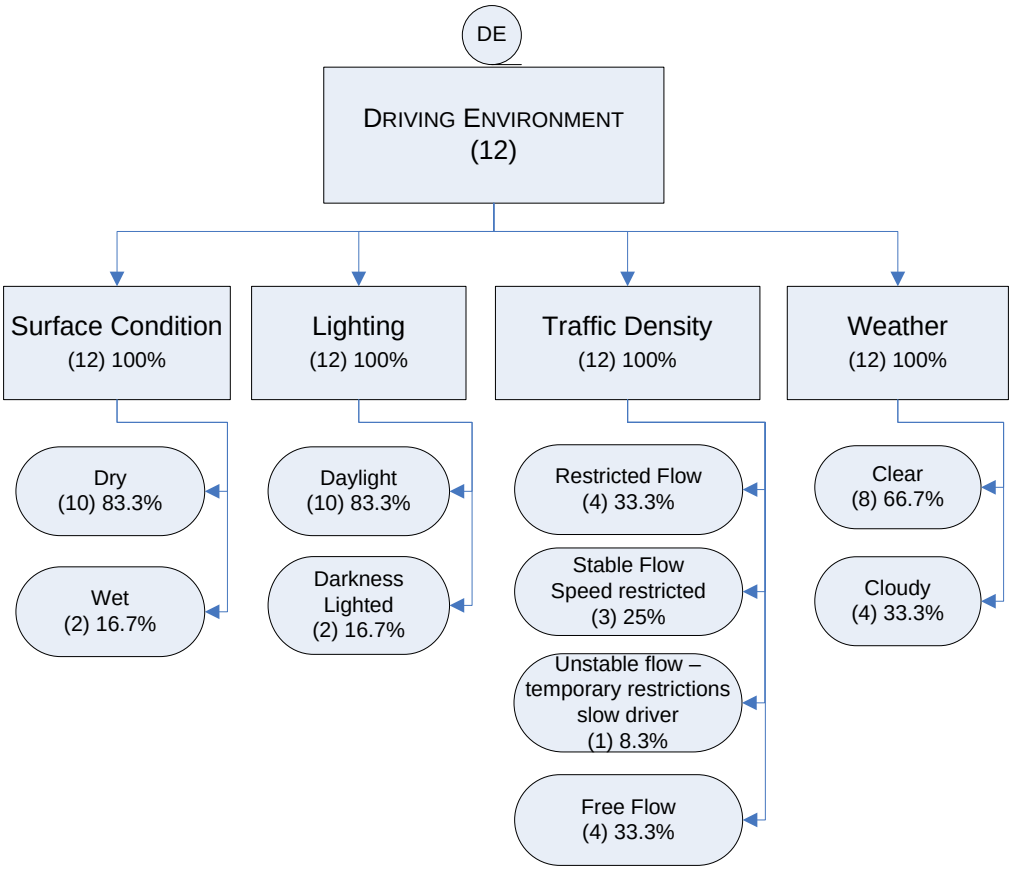


Figure 5.26. Breakdown of driving environment variables for crashes involving following vehicles.

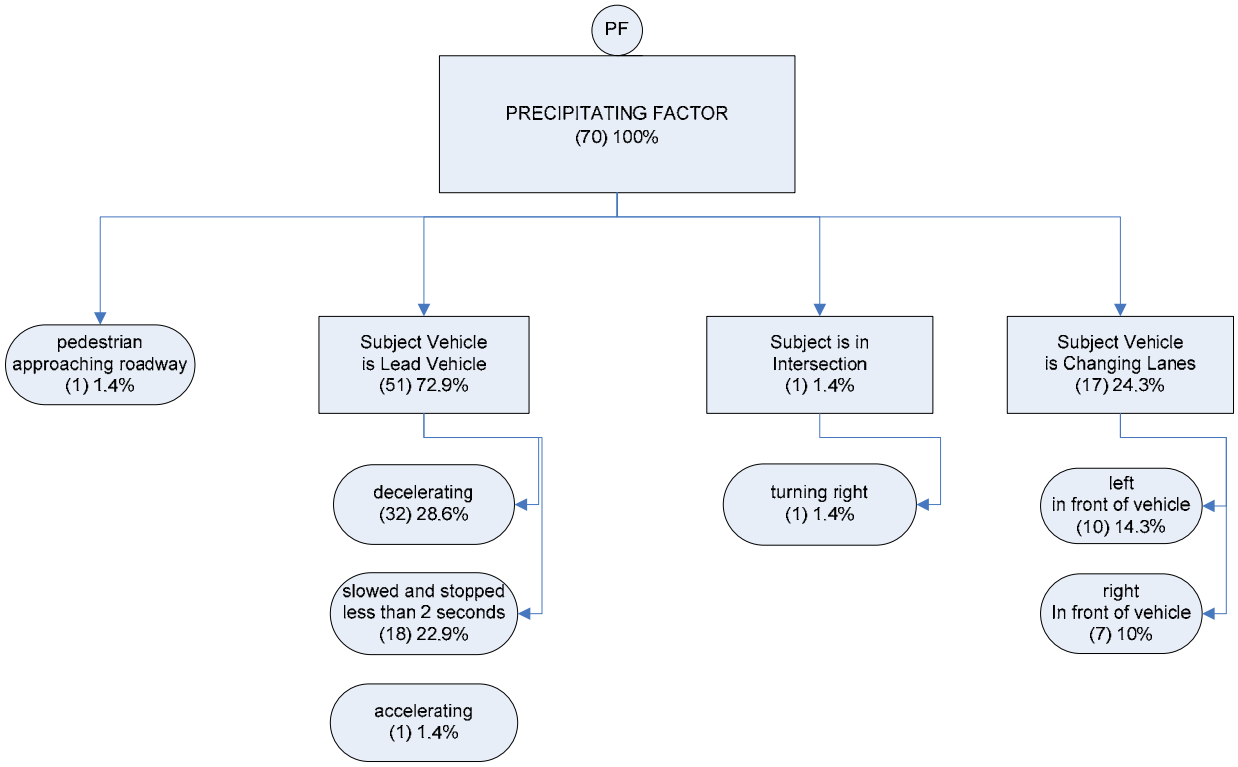


Figure 5.27. Breakdown for precipitating factors for near-crashes involving following vehicles.

Driving proficiency was identified as a contributing factor in more of the near-crashes (49%) than the crashes (33%). Aggressive driving (17%) and drowsiness (13%) were also contributing factors in these following-vehicle near-crashes.

Thirty percent of the drivers in the following-vehicle near-crashes had no reaction. Although the majority of the drivers braked alone (49%), some drivers braked and steered (13%), steered alone (3%), or accelerated (4%). Interestingly, over 8 percent of the driver avoidance maneuvers included steering and/or accelerating to avoid the conflict.

Although weather was not a contributing factor, it was an associated factor with over 12 percent of the near-crashes, including inclement weather and wet surface conditions (Figure 5.28). Only 23 percent of the near-crashes were associated with a free-flow of traffic. Roadway alignment may have played less of a role with near-crashes than with crashes in this conflict type, with 11 percent of the near-crashes being on curves as compared to 42 percent of the crashes being on curves. Twenty-seven percent of the crashes were intersection-related (Figure 5.29 infrastructure).

When looking for other contributing factors, visual obstructions were present in 6 crashes. Three were sunlight glare, one was a moving vehicle, and one was road sight distance due to a curve or hill.

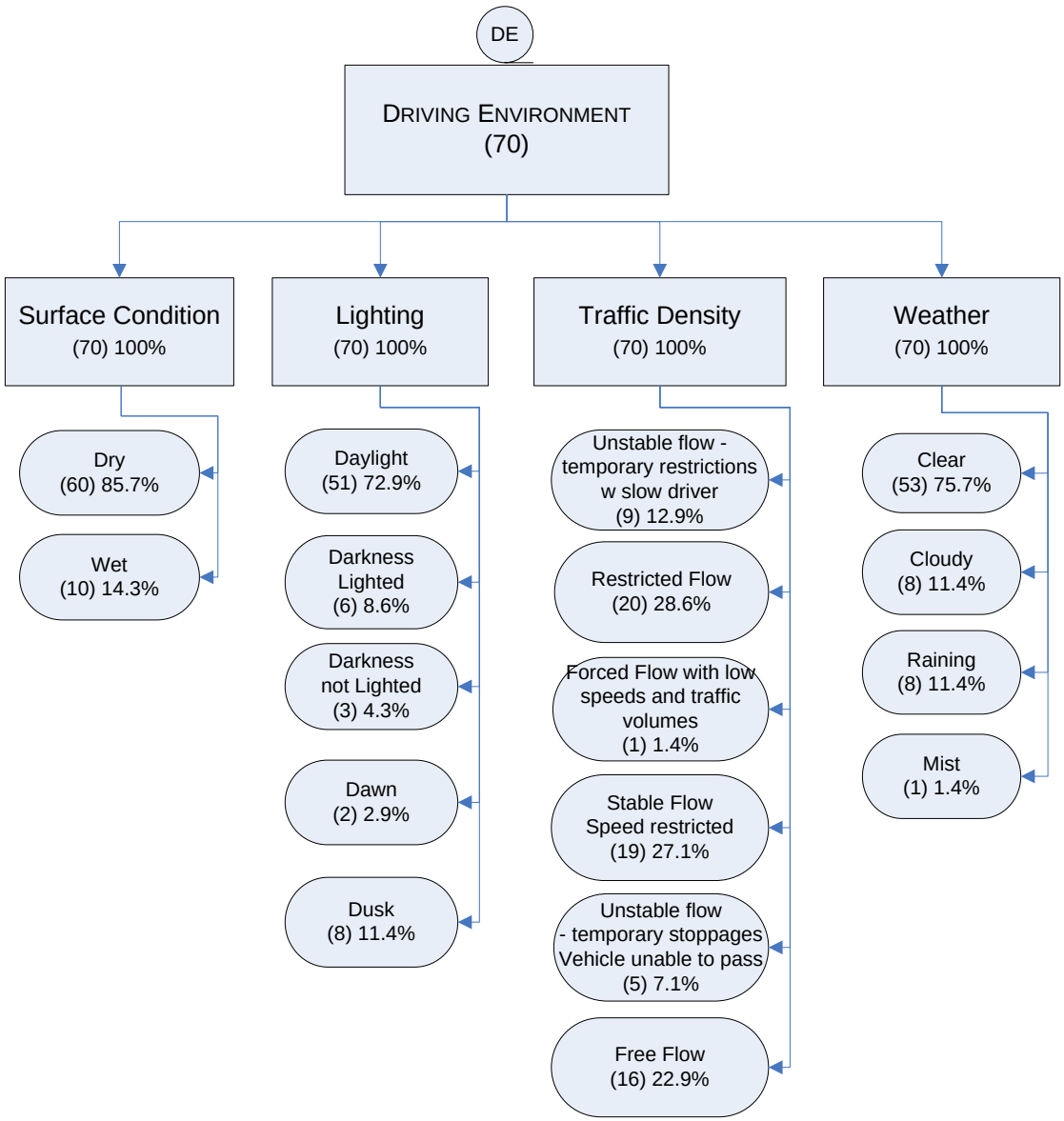


Figure 5.28. Breakdown of driving environment variables for near-crashes involving following vehicles.

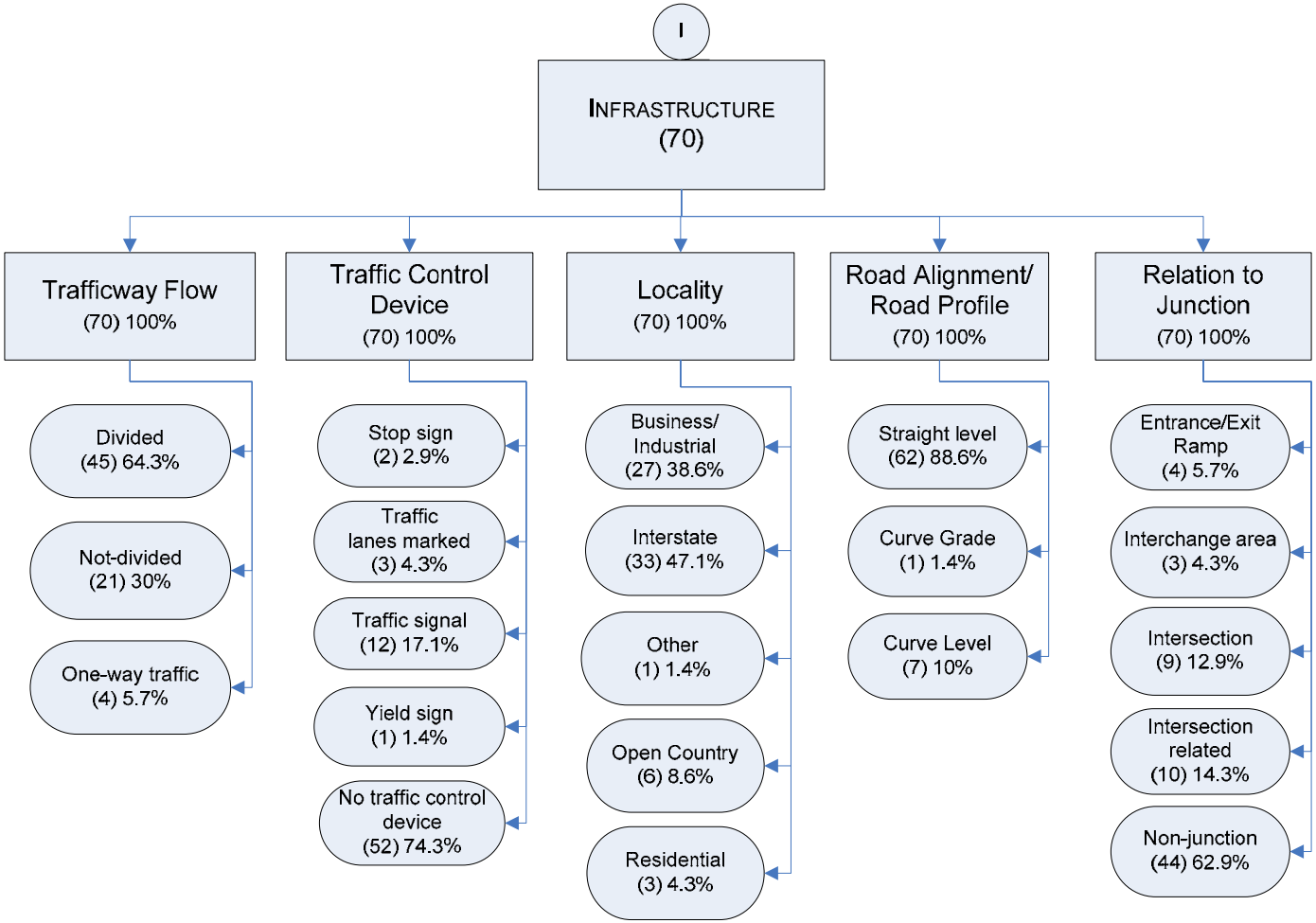


Figure 5.29. Breakdown of infrastructure-related variables for near-crashes involving following vehicles.

Following Vehicle Incidents. More varied precipitating factors were found for the 776 following-vehicle incidents than for the following-vehicle crashes or near-crashes (Figure 5.30). Although no following-vehicle crashes and 24 percent of the near following-vehicle crashes were associated with the subject vehicle changing lanes, 46 percent of the incidents included the subject vehicle moving into the traffic lane of the conflict. As with the near-crashes, the maneuver was more common to the left (27%) than the right (18%) for the incidents. In 31 percent of the incidents, the subject vehicle was decelerating. In 10 percent of the incidents, the subject vehicle was stopped for more than 2 seconds. An additional 7 percent were stopped for less than 2 seconds as the precipitating factor for the following-vehicle incidents.

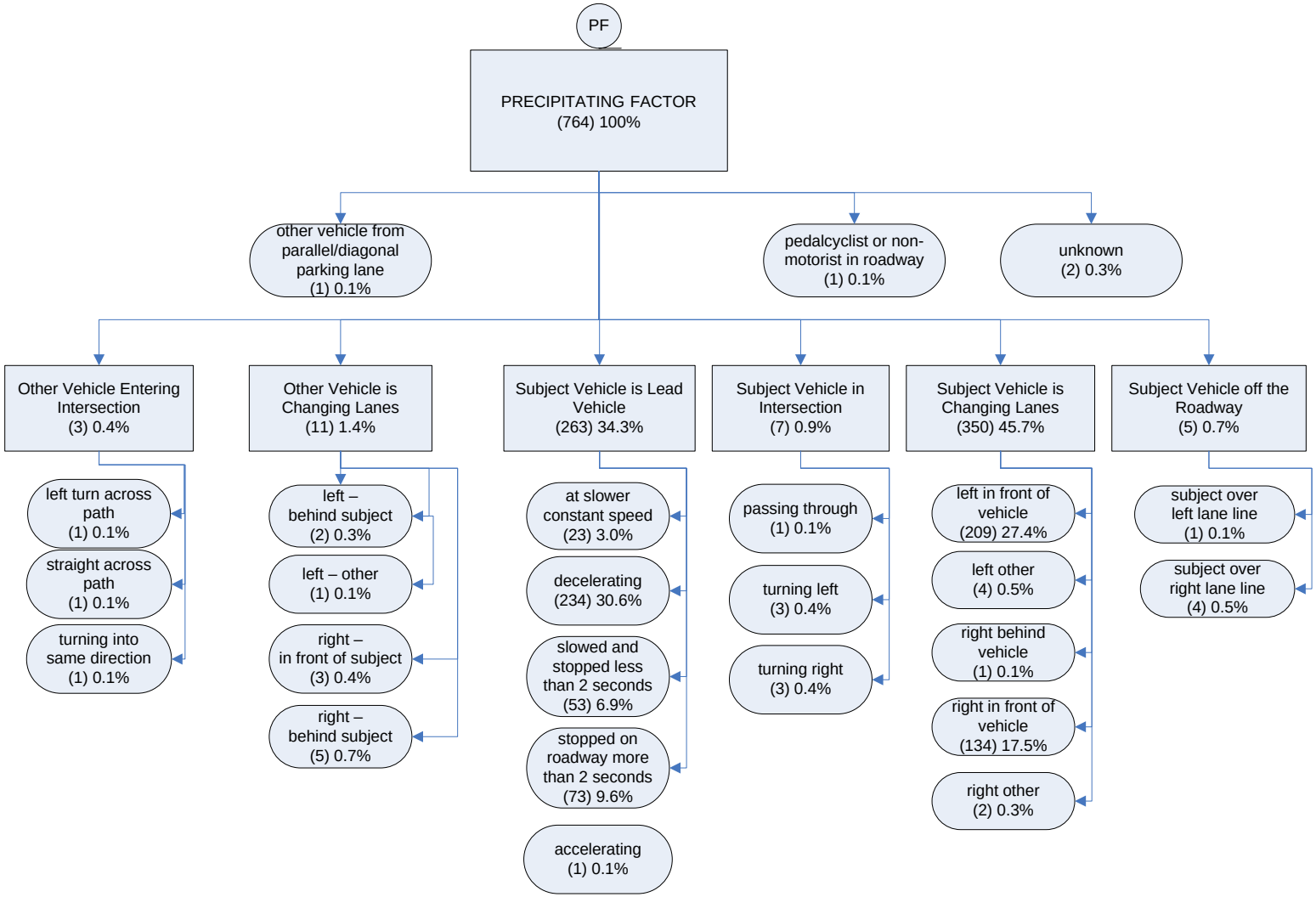


Figure 5.30. Breakdown for precipitating factors for incidents involving following vehicles.

Forty-six percent of the drivers in the following-vehicle incident had no reaction for the avoidance maneuver. As with the near-crashes the majority of the incident avoidance maneuvers were braking alone (38%); some drivers braked and steered (5%), steered alone (4%), or accelerated (7%).

The only other contributing factors that were greater than 1 percent were 33 glare incidents due to sunlight (4%) and 9 visibility decrement incidents (1%) due to things such as rain, snow, dust, etc. The frequency of the other contributing factors leading to a visibility decrement is as follows: inadequate roadway lighting (1); moving vehicle (1); road infrastructure (1); and other obstructions (2). Roadway delineation was a contributing factor in 5 of the incidents, and roadway alignment was a contributing factor in three incidents.

Although inclement weather was a contributing factor in approximately one percent of the incidents, it was an associated factor in 5 percent of the incidents, and wet surface conditions were associated factors in 7 percent (Figure 5.31). Daylight was associated with 75 percent of the incidents, and only 11 percent were in free flow traffic conditions.

Less than 6 percent of the incidents were on curves as compared to 42 percent of the crashes being on curves. Twenty-four percent of the incidents were intersection-related (Figure 5.32).

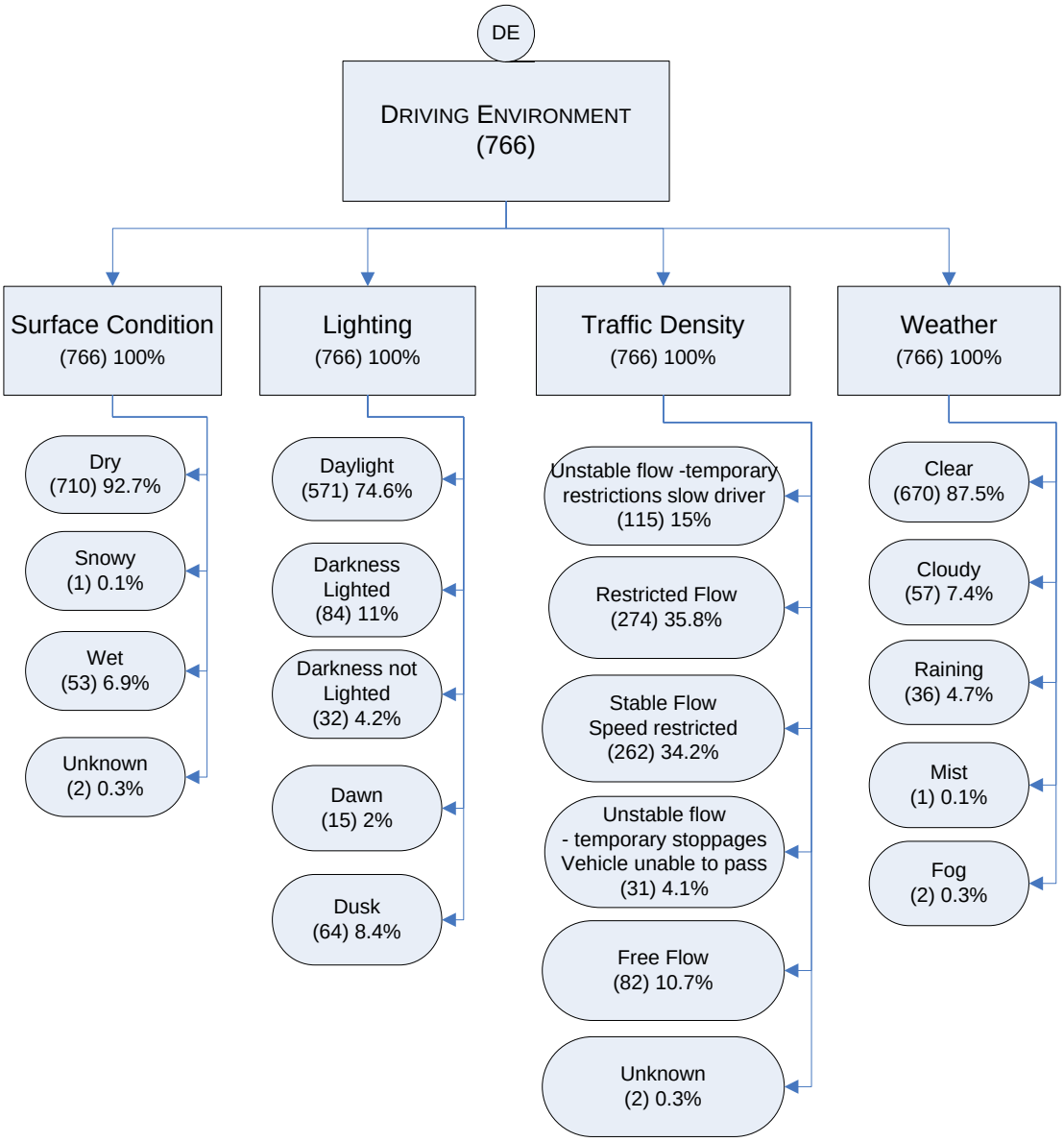


Figure 5.31. Breakdown of driving environment variables for incidents involving following vehicles.

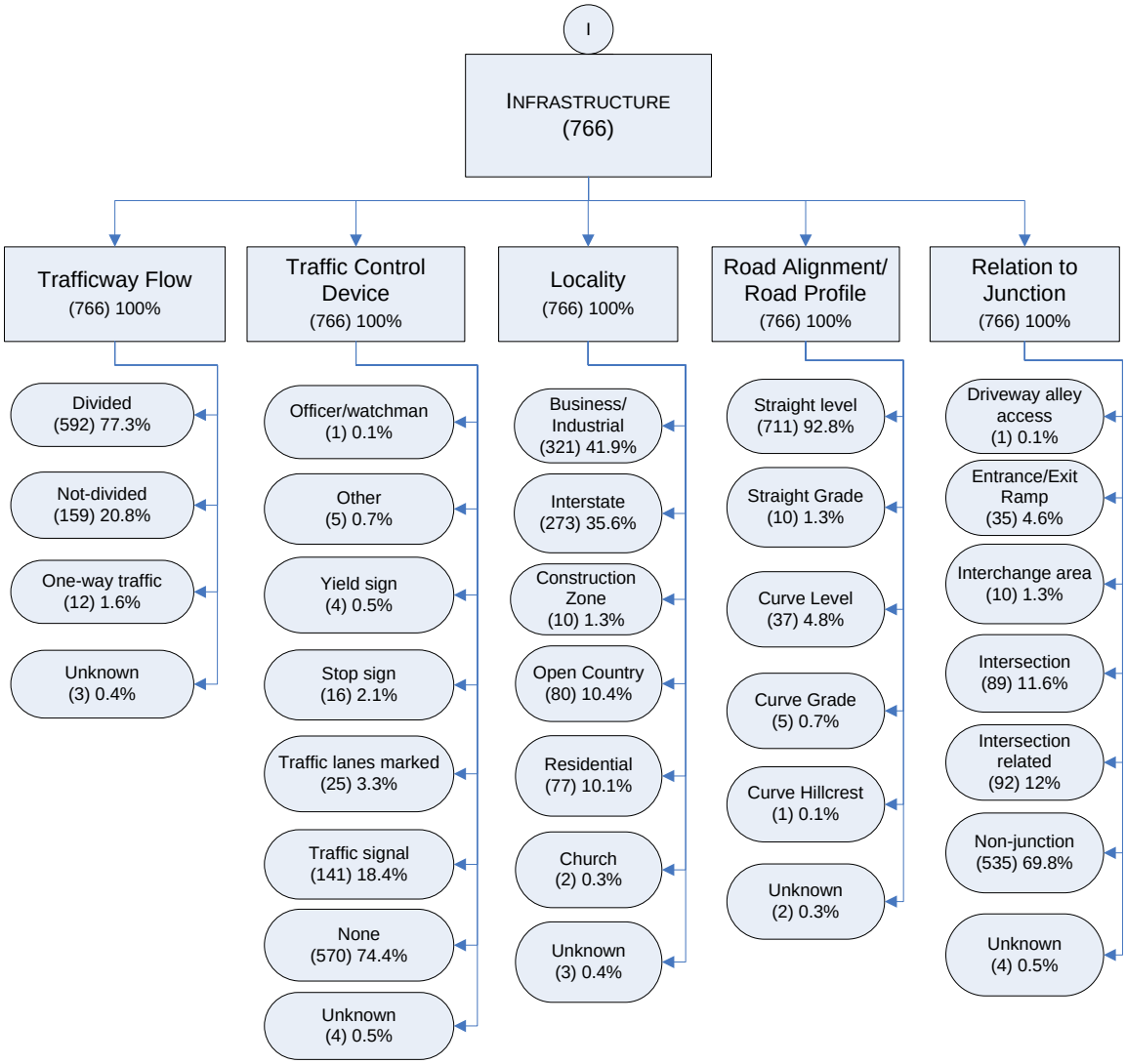


Figure 5.32. Breakdown of infrastructure variables for incidents involving following vehicles.

Object/Obstacle Conflicts
Object or obstacle conflicts involve events in which there is an object or obstacle in the lane of travel that subjects must respond to in order in the attempt to avoid a crash. An inappropriate response or failure to respond resulted in a crash or collision with obstacle. Object/obstacle conflicts accounted for 12 percent of the crashes, 1 percent of the near-crashes, and 5 percent of the incidents. Of the 9 crashes, 7 were classified as other, one was backing into a fixed object, and the last one included departing the road. The crashes classified as other included parking gate, debris flying in roadway, and so forth.

Object/Obstacle Crashes. [diagrams cover multiple object sections] The subject vehicle in three out of the 9 obstacle crashes (33%) was going straight at a constant speed. Two were decelerating in the traffic lane, two were turning right, one was making a u-turn, and the final one was backing but not parking.

For the precipitating factor in 5 of the 9 crashes (56%), the obstacle was in the road. Two of the 9 (22%) crashes had excessive speed as a contributing factor. One of the crashes had the subject over the right lane line. The final crash was an end departure as the precipitating factor, such as driving through the dead end portion of a roadway.

Inattention to the forward road was a contributing factor in 5 of the 9 crashes (55%). Two of the crashes included driving-related inattention, with drivers looking out the left window (22%). Two of the crashes included drivers talking on the cell phone (22%), and the final crash included drivers interacting with an object in the vehicle. In other driver factors, driver proficiency was a contributing factor in 4 of the 9 of the crashes (44%). Three of the crashes had aggressive driving as the contributing factor (33%), and one of these crashes was classified as drowsiness-related.

Not surprisingly, 7 of the 9 drivers did brake prior to crashing as an avoidance maneuver. Two of the 9 combined braking with steering to the left. Only one of the 9 had no reaction to the obstacle.

Three of the 9 crashes had an infrastructure contributing factor. One was roadway alignment, one was roadway delineation, and the other was related to traffic control device. Visual obstructions were present in two of the 9 crashes. A moving vehicle and trees/crops/vegetation were the two obstructions cited.

For the associated factors only 2 of the 9 were during daylight (Figure 5.33). Weather was not an associated factor, and surface condition was only an associated factor in one of the crashes (snowy surface condition). The associated infrastructure included a curve in two of the crashes, and three of these crashes were in parking lots (Figure 5.34).

Object/Obstacle Near-Crashes. For all drivers, the pre-incident maneuver of the obstacle near-crashes included presence in the traffic. Five vehicles were going straight at a constant speed, and one was decelerating.

For the precipitating factor in 5 of the 6 obstacle near-crashes (83%), the obstacle was in the road.

When considering driver-associated factors, 2 of the 6 near-crashes had aggressive driving as a contributing factor, and one had driver proficiency as a contributing factor.

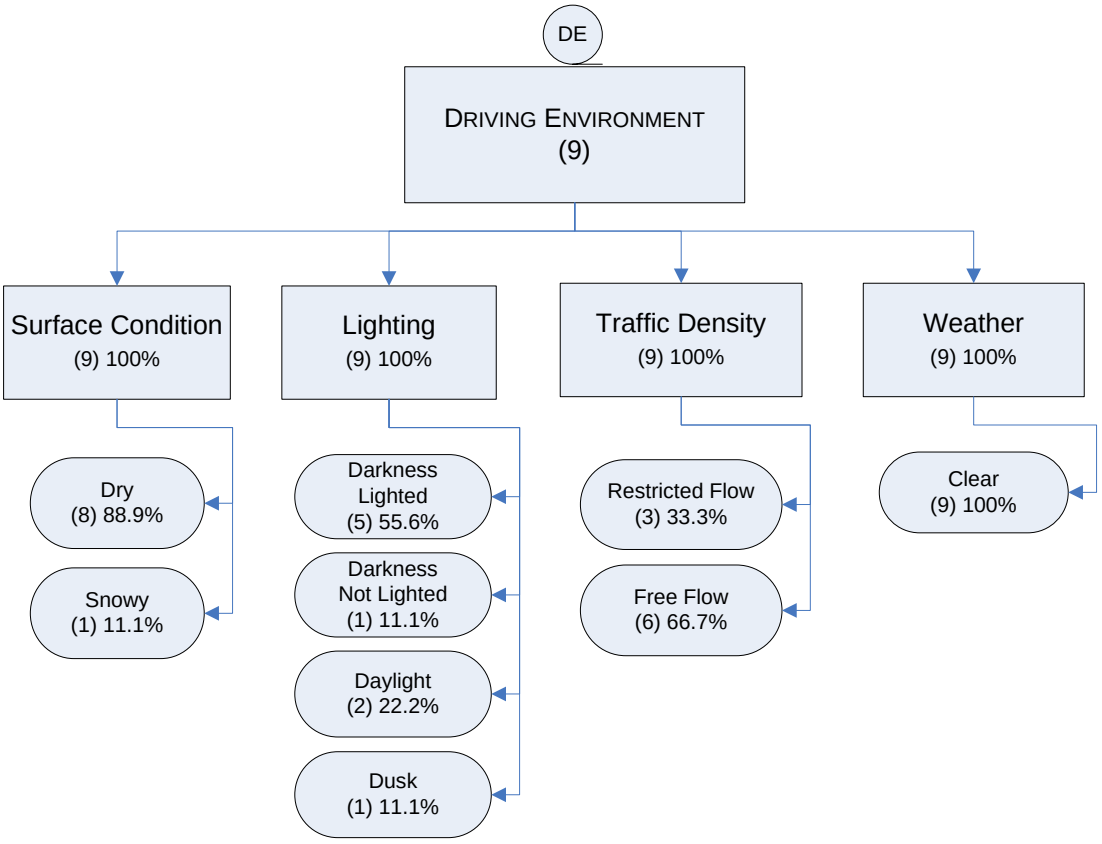


Figure 5.33. Breakdown of driving environment variables involving crashes with obstacles/objects.

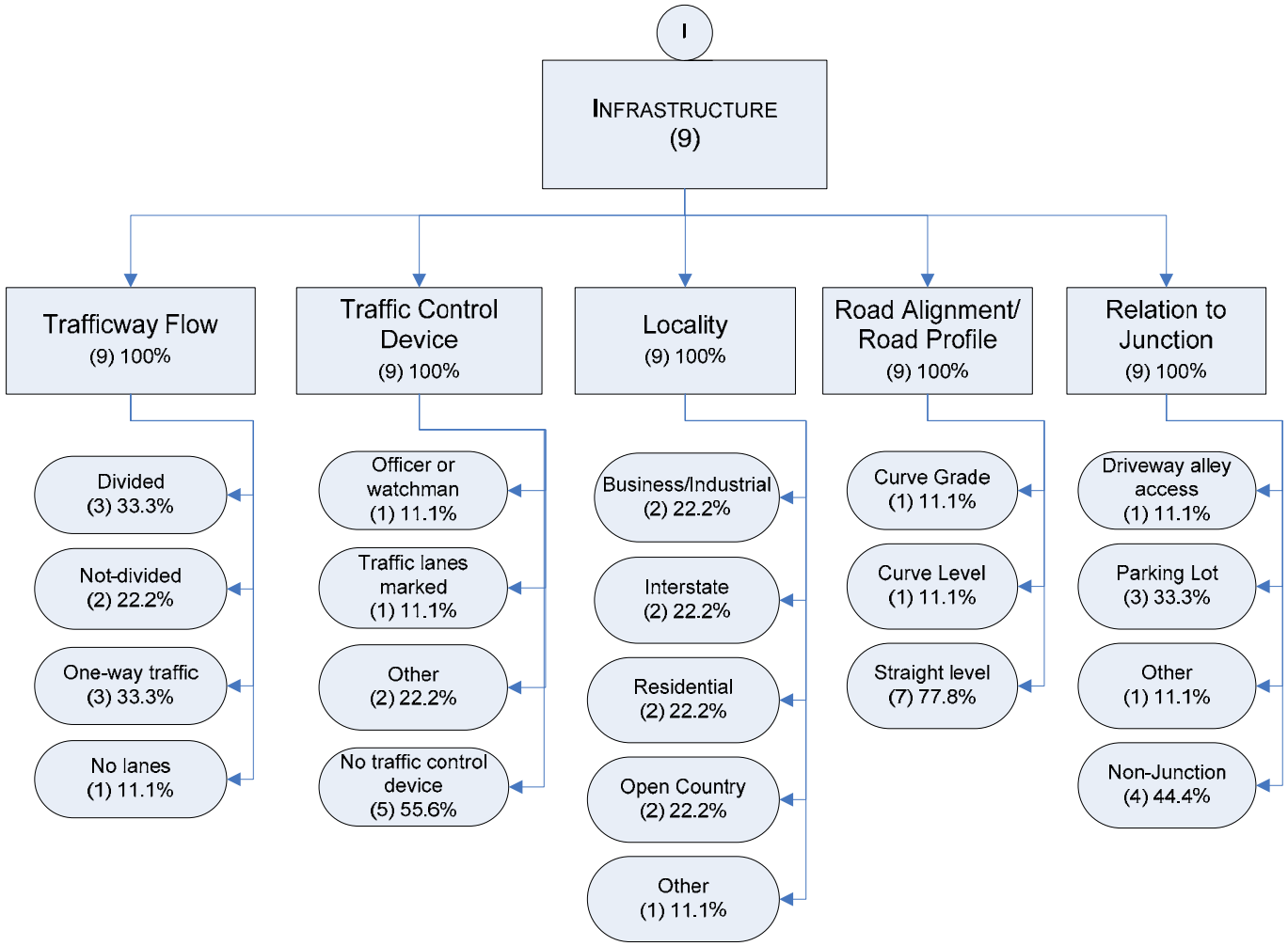


Figure 5.34. Breakdown of infrastructure-related variables for crashes involving objects or obstacles.

Inattention to the forward road was a contributing factor in 2 of the 6 near-crashes. In one near-crash the driver was dining, and in the other near-crash the driver had a passenger in the adjacent seat.

All 6 drivers braked as the avoidance maneuver. Three drivers also steered right and 2 also steered left.

The other contributing factors indicated that roadway alignment, roadway delineation, and glare due to sunlight were each associated with a near-crash.

As for the associated factors all the near-crashes were associated with clear weather, dry surface condition, straight roadway alignment, and were non-intersection-related. Only 1 of the 6 crashes was in free-flow conditions.

Object/Obstacle Incidents. There were 394 obstacle-related incidents identified. Ten percent of the objects were in an unknown location due to video resolution or nighttime conditions. Prior to the crash, the vast majority of the subject vehicles were going straight (91%). The pre-incident maneuver included going straight at a constant speed (50%), going straight and accelerating (22%), and decelerating in the traffic lane (19%).

For the precipitating factor in 70 percent of the incidents, the obstacle was in the road. Poor road conditions were the precipitating factor in 18 percent of the incidents.

Inattention to the forward road was a contributing factor in 16 percent of the incidents. Two percent of the incidents included driving-related inattention, with drivers looking out the left window (3 incidents) and the right window (3 incidents). The remaining driver-related inattention included looking at the center mirror. As shown in Figure 5.35, the remaining 14 percent were due to secondary task distraction, with the biggest contributing factors being cell phone use (5%) and a passenger in the vehicle (5%). As in the other conflicts, talking or listening accounted for the majority of the cell phone-related inattention (4%). As for the other driving-related factors, driver proficiency was a contributing factor in 20 percent of the incidents. Only 2 percent of the incidents were aggressive-driving-related, and 5 percent of the incidents were drowsiness-related.

Steering alone was a surprisingly common avoidance maneuver. Twenty-seven percent steered to the left, and 19 percent steered to the right. An additional 14 percent combined braking with steering to the left and 10 percent combined braking with steering to the right. Another 16 percent only braked to avoid the obstacle, and 11 percent had no reaction.

Infrastructure played a contributing role in approximately one third of the incidents. Roadway alignment (7%), roadway delineation (25%), and traffic control devices (1%) each contributed to these incidents (Figure 5.36). Reduced visibility due to rain or snow was a contributing factor in three of the incidents (1%). Glare due to sunlight was a contributing factor in 7 of the incident (2%), and glare due to headlights contributed in

one incident. Three incidents were classified as including inadequate roadway lighting, and two incidents had visual obstruction due to trees/crops/vegetation.

Also shown in Figure 5.36, inclement weather was an associated factor in 8 percent of the incidents, and non-dry surface conditions were an associated factor in 17 percent of the incidents. One third of the incidents were during non-daylight, and one third were in non-free flow traffic conditions. For fixed infrastructure, 11 percent of the incidents were on curves, and 10 percent of the incidents were intersection-related (Figure 5.37).

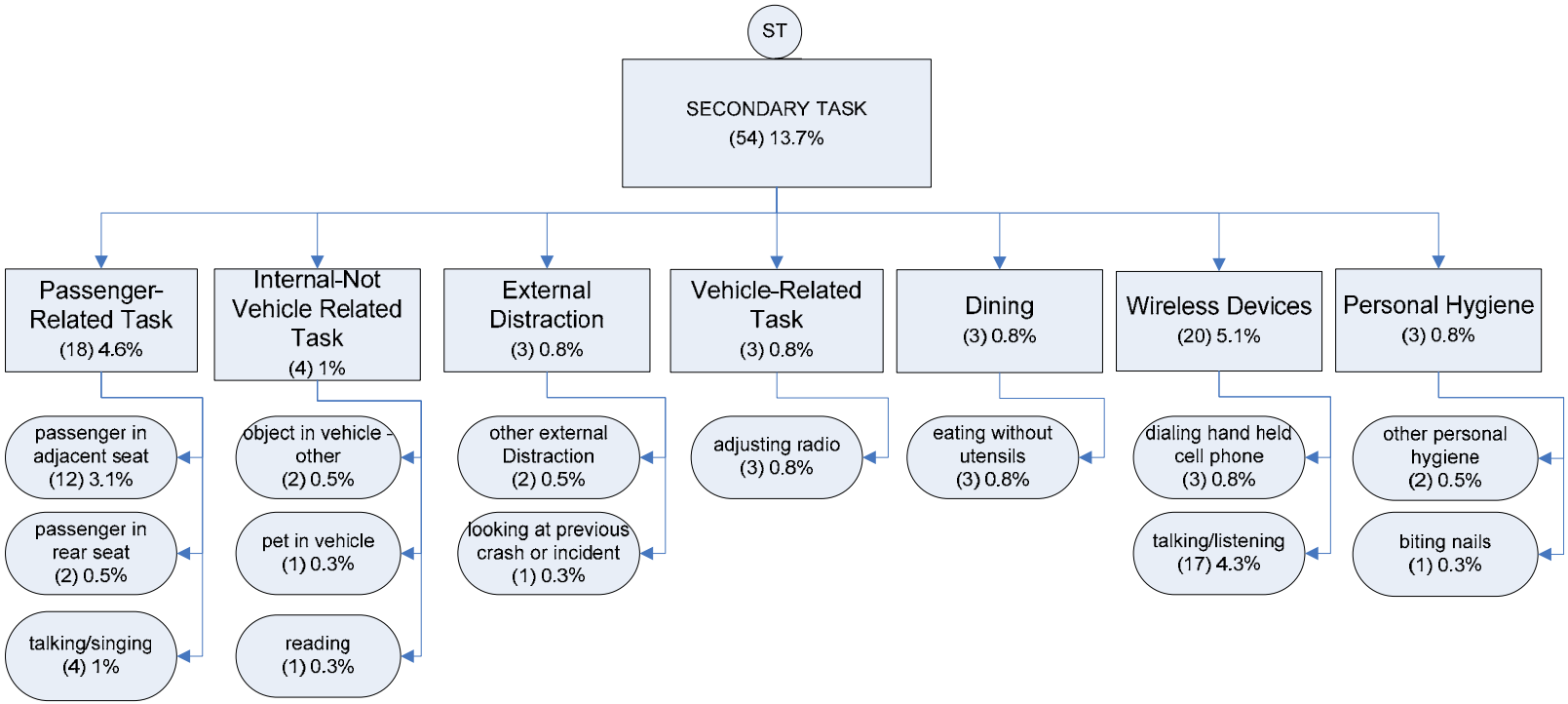


Figure 5.35. Breakdown of secondary tasks for incidents involving obstacles or objects.

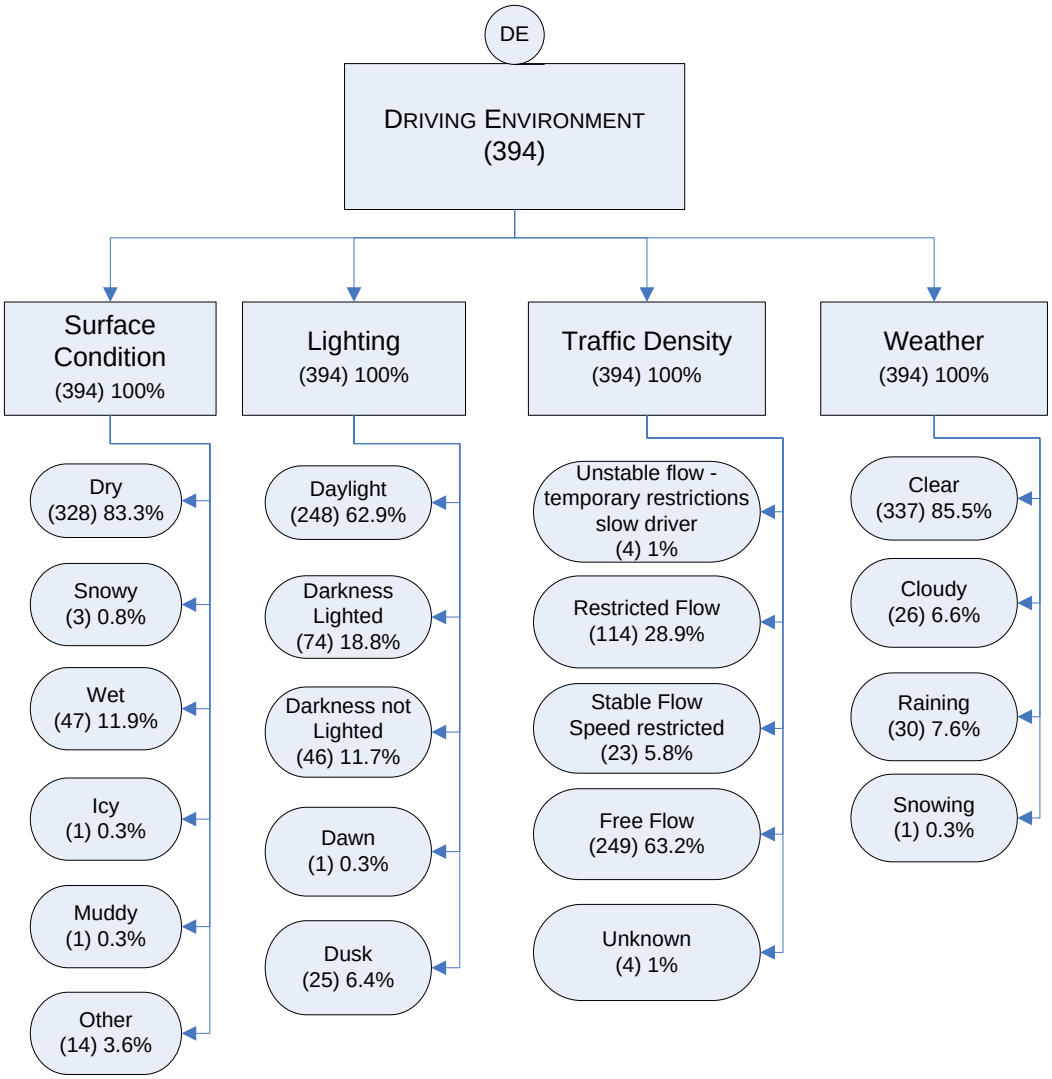


Figure 5.36. Breakdown of driving environment variables for incidents involving objects or obstacles.

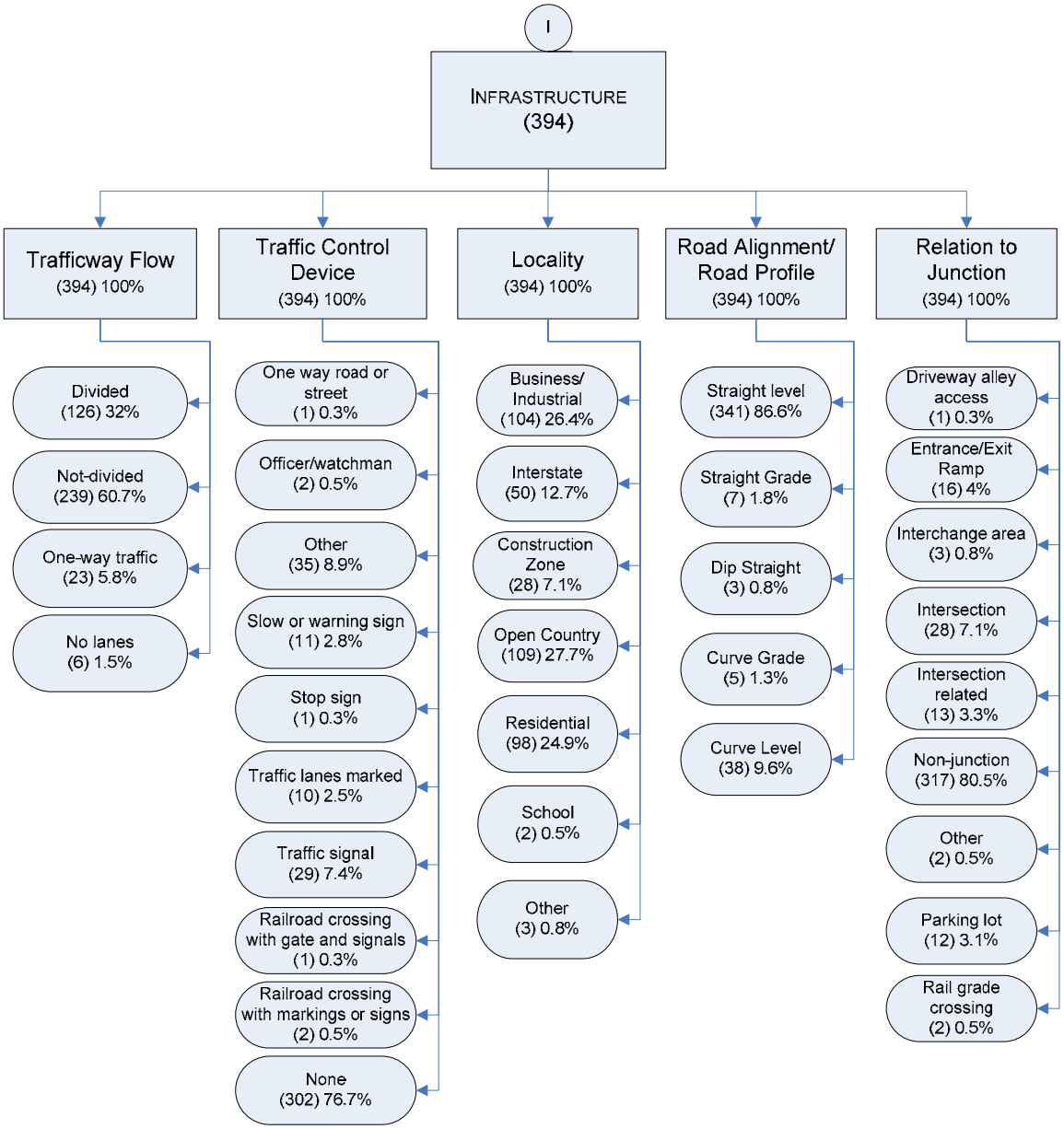


Figure 5.37. Breakdown of infrastructure-related variables for incidents involving obstacles or objects.

Animal Conflict
Animal conflicts accounted for 3 percent of the crashes, 1 percent of the near-crashes, and 1 percent of the incidents.

Animal Crashes. Prior to both crashes the driver was going straight at a constant speed. Not surprisingly, the precipitating factor in the two animal crashes was an animal in the road. Inattention to the forward roadway, drowsiness, and aggressive driving did not contribute to these crashes. During both crashes the driver braked alone to attempt to avoid the animal. Infrastructure and driving environment did not contribute to either. There were no relevant associated weather conditions, surface conditions, and roadway alignment to these crashes either. This lack of contributing factors is expected with this type of crash.

Animal Near-Crashes. Prior to the near-crashes the driver was going straight at a constant speed (7), going straight accelerating (1), negotiating a curve (1), or turning left (1). Not surprisingly, the precipitating factor in the 10 animal near-crashes was also an animal in the road in 9 of the near-crashes. In the remaining near-crash the animal was approaching the roadway. Inattention to the forward roadway was a contributing factor in three of the near-crashes. In two near-crashes, a passenger was in the adjacent seat, and in the other near-crash, the driver was talking/listening on the cell phone. As with the crashes, aggressive driving did not contribute to these near-crashes. However, drowsiness was a contributing factor in 4 of the near-crashes. During all of the near-crashes, the driver braked alone to avoid the animal. The only other contributing factor was in one crash when there was limited sight distance due to a hill or obstruction.

For the associated factors, 9 of the 10 near-crashes were in darkness, but 3 were lighted. Only clear weather and dry surface conditions were associated with these near-crashes. Three of these near-crashes were on a curve.

Animal Incidents. Of the 56 animal incidents, 75 percent had animals in the road as the precipitating factor. In 23 percent of the incidents, the animal was approaching the roadway. Prior to the incidents, the driver was going straight at a constant speed (66%), going straight accelerating (20%), negotiating a curve (7%), changing lanes (4%), or turning right (2%).

Inattention to the forward roadway was a contributing factor in 5 of the incidents (9%). The 5 incidents had 5 different secondary task distractions, including a passenger in the adjacent seat, reaching for an object, talking/listening on the cell phone, eating with utensils, and an external distraction. As with the crashes and near-crashes aggressive driving did not contribute to these incidents. However, drowsiness was a contributing factor in 13 percent of the incidents, and 5 percent of the incidents had driver proficiency errors.

Unlike the crashes and the near-crashes, drivers avoided these incidents by steering. In 13 percent of the incidents, drivers steered to the left, and in 4 percent of drivers steered to the right to avoid the animal. In an additional 13 percent of the incidents, drivers braked and steered to

the left, and in 7 percent, drivers braked and steered to the right to avoid the animal. In the remaining 64 percent of the incidents, drivers braked alone to avoid the animal.

Only two of the 56 incidents had any other contributing factor. One was limited sight distance due to curve or hill. The other was glare due to sunlight.

For the associated factors, many of the incidents were in darkness (63%) and lit darkness (16%). Inclement weather and surface conditions were present in only two of the incidents. Eight percent of the incidents were in free flow traffic conditions. As for road infrastructure, 20 percent of the incidents were on a curve, and one was intersection-related.

Parked Vehicle Conflicts

Parked vehicle conflicts accounted for 6 percent of the crashes, 1 percent of the near-crashes, and 1 percent of the incidents. Two of the 4 crashes were backing into a fixed object. One was backing into traffic, and the other was being sideswiped in the same direction.

Parked Vehicle Crashes. Four different precipitating factors were associated with the 4 crashes. One crash had participant over the left road edge, another had the subject vehicle attempting a lane change, another was backing from the driveway, and the remaining one was an end departure. In two crashes, the driver braked alone attempting to avoid the crash. In one crash, there was no reaction, and the remaining crash was classified as "other."

Inattention to the forward roadway was a contributing factor in 2 crashes. In one crash, the inattention was a cognitive distraction, and in the other, it was a passenger in the adjacent seat. For the other driver factors, one of the crashes had aggressive driving as a contributing factor, and one had driver proficiency as a contributing factor. Drowsiness did not contribute to these crashes.

Roadway delineation contributed to one crash, and a parked vehicle provided a visual obstruction for another crash. No other contributing factors were present.

Inclement weather, wet surface conditions, and curved roadway alignment were not associated with these crashes either. This lack of contributing factors is expected with this type of crash.

Parked Vehicle Near-Crashes. Four different precipitating factors were associated with the 5 near-crashes. Two of the near-crashes involved another vehicle backing, and one had another vehicle leaving a parallel diagonal parking lane. One of the other near-crashes had a lead vehicle stopped in the roadway more than 2 seconds, and the remaining near-crash involved a pedestrian approaching the roadway.

Aggressive driving, driver proficiency errors, and drowsiness were not contributing factors. However, inattention to the forward roadway was a factor in three of the 5 near-crashes. There were two driving-related inattention contributing factors. One of the factors was the left mirror; the other was the right mirror. The remaining inattention contributing factor was looking at a pedestrian. There were no other contributing factors present.

Braking was the most common avoidance maneuver. In two near-crashes the driver braked alone, and in another two near-crashes the driver braked and steered left to avoid the crash. In the remaining near-crash the driver steered left without braking to avoid the crash.

As with crashes, inclement weather and surface conditions were not associated with these near-crashes. Two of the near-crashes occurred in a parking lot.

Parked Vehicle Incidents. There were 83 incidents found for parked vehicle conflicts. Of those 83 incidents, a lead vehicle stopped on the roadway more than 2 seconds was the largest precipitating factor (46%) (Figure 5.38). These events were caused by vehicles either stopped or parked in a travel lane. Another vehicle pulling out from a parallel diagonal parking lane was the second biggest precipitating factor (17%). The subject vehicle off the roadway was the next biggest contributing factor (14%).

Although these driver factors, other than driver inattention, were not identified as contributing to the near-crashes, aggressive driving (16%), driver proficiency errors (39%), and drowsiness (5%) were contributing factors in the parking incidents. Inattention to the forward roadway was also a contributing factor (25%). There were three driving-related inattention contributing factors. One of the factors was the left mirror, the other two were the right mirror. Secondary tasks were a contributing factor in 22 percent of the incidents, with cell phone use accounting for half the contributing factors (Figure 5.39).

Steering as an avoidance maneuver was the most common, with over three-quarters of the incidents including a steering maneuver. Steering to the left (36%) and steering to the left and braking (23%) were the largest avoidance maneuvers. Steering to the right (7%) and steering to the right and braking (6%) were also avoidance maneuvers. Only 18 percent used braking alone as the avoidance maneuver. An additional 5 percent accelerated and steered, and 2 percent had no reaction.

Roadway alignment (6%), roadway delineation (8%), and roadway sight distance (1%) were all identified as contributing factors. In addition to the roadway sight distance, there were visual decrements due to rain, snow, or fog (2%), inadequate roadway lighting (1%), glare due to sunlight (2%), and trees, crops, or vegetation (1%).

For the associated factors wet or snowy surface conditions (10%), rain (2%), and restricted flow were each associated with 57 percent of the incidents. Curved roadway (15%) and intersection-related (6%) were also associated factors.

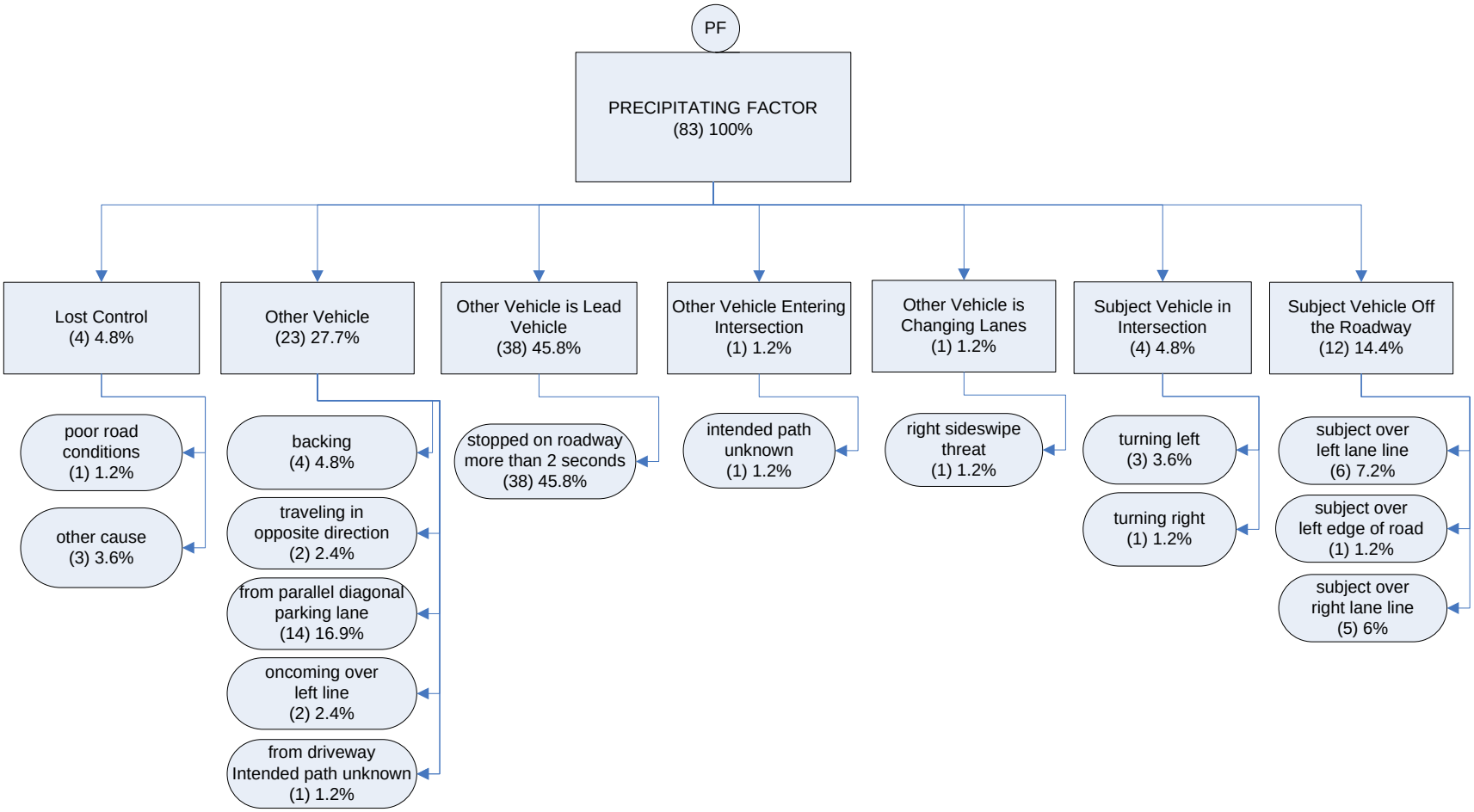


Figure 5.38. Breakdown of precipitating factors for incidents involving a parked vehicle.

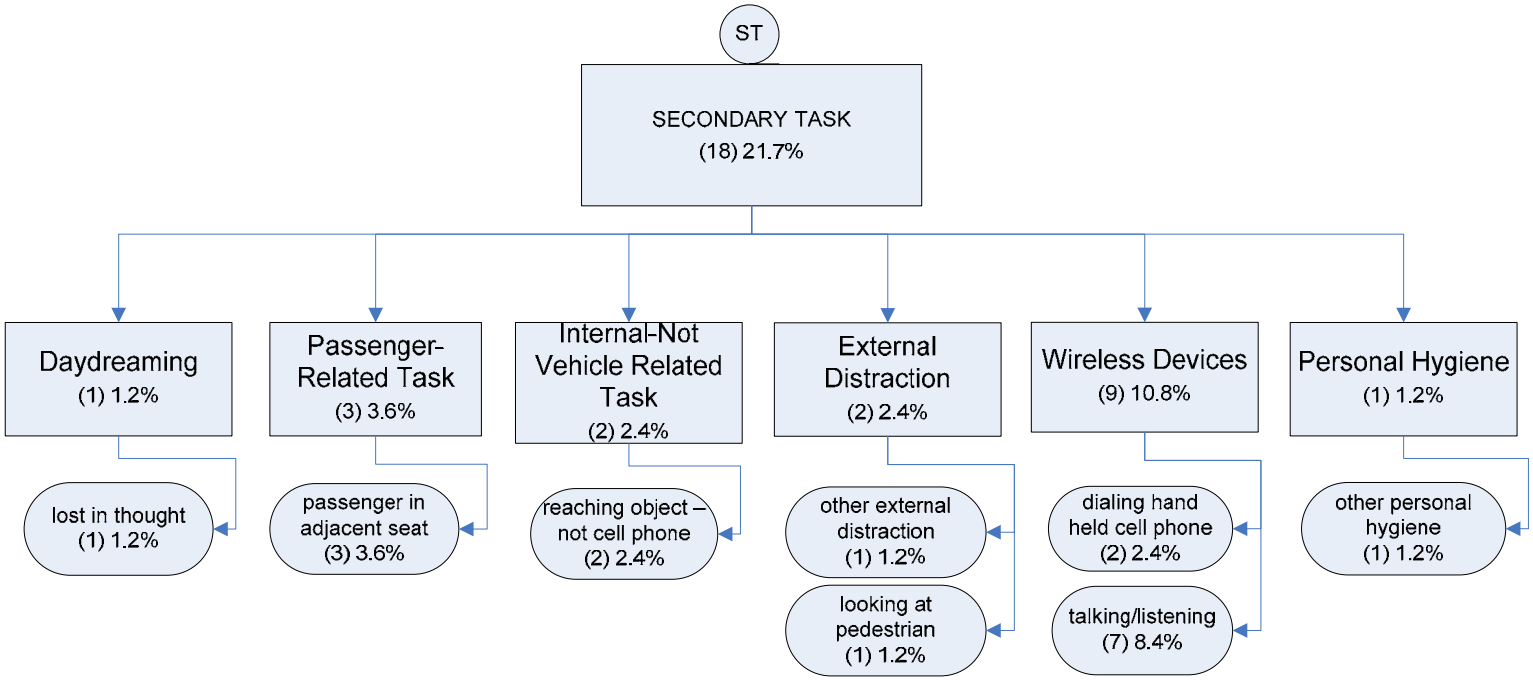


Figure 5.39. Breakdown of secondary tasks for incidents involving a parked vehicle.

Conflict with Vehicle Turning Across Subject Vehicle Path in Opposite Direction
A vehicle turning across subject vehicle path in opposite direction accounted for 3 percent of the crashes, 4 percent of the near-crashes, and 1 percent of the incidents.

Crashes with Other Vehicle Turning Across Subject Vehicle Path in Opposite Direction. Of the two LTAP/OD crashes, one occurred on a congested roadway with offset intersection approaches (no signal) and the other occurred at a signalized intersection when the other driver made a left turn on a red arrow.

The precipitating factor for the crash that occurred at a signalized intersection, was that the other vehicle entered the intersection making a left turn across subject vehicle's path. This driver attempted to brake without locking up the wheels to avoid the crash. This crash also occurred in the rain on a lit roadway at night. The traffic congestion was labeled free flow.

The precipitating factor for the crash that occurred at an offset intersection was subject vehicle attempting a left turn. This attempt occurred after the other vehicle was approaching (head-on crash). There was no reaction by this driver.

Considering the driver factors, there was no aggressive driving, no driver impairments, and no driver proficiency errors. However, both drivers were not attentive to the forward roadway. One driver had a passenger in vehicle, and the other was looking out the right window.

This crash occurred during daylight hours in stable traffic flow. Roadway alignment was identified as a contributing factor for this crash, but not surprisingly there were no visual obstructions.

Near-Crashes with Other Vehicle Turning Across Subject Vehicle Path in Opposite Direction. Out of 27 near-crashes, 74 percent involved a vehicle turning left across the subject vehicle's path, and 19 percent involved a vehicle turning the opposite direction to that of subject vehicle's path. Only 4 percent of near-crashes involved the subject vehicle passing through other vehicle's path at an intersection. An additional 4 percent involved the subject vehicle doing a lane change.

For the driver factors, 26 percent of the near-crashes were classified as driver proficiency errors. An additional 11 percent of near-crashes were classified as aggressive driving, and 11 percent were classified as drowsiness-related. Inattention to forward roadway was a contributing factor in 37 percent of the near-crashes. These factors included adjusting the radio (7%), passenger in vehicle (4%), and cell phone talking or listening (4%). Seven percent of near-crashes were due to driving-related inattention when drivers were checking their left or right side-view mirrors.

Because the majority of near-crashes occurred when subject vehicles were going straight, the drivers commonly reacted by braking with no lockup (44%); other times, drivers mostly responded by steering either left or right and combining braking and steering both as avoidance maneuvers.

Visual decrements were a contributing factor in over 30 percent of the near-crashes. These factors included glare due to sunlight (4%), moving vehicles (11%), and parked vehicles (7%). Road sight distance was also a contributing factor in one near-crash.

The associated factors for near-crashes, as with crashes, were primarily clear weather (82%) under daylight (63%) conditions. Eighty-six percent of near-crashes were intersection-related. Nearly three-fourths of the near-crashes occurred in a business or industrial locality, while only 4 percent occurred on interstate roadways and 11 percent in residential areas. Eighty-six percent of the time, the road alignment was straight. The surface condition of roads was dry during 82 percent of near-crashes. Thirty-three percent of near-crashes were due to restricted traffic flow, while 44 percent occurred in free flow conditions. Moreover, 78 percent of near-crashes occurred in non-divided traffic flow conditions.

Incidents with Other Vehicle Turning Across Subject Vehicle Path in Opposite Direction. Out of 79 incidents, not surprisingly the biggest precipitating factors occurred when other vehicles were entering an intersection (81%) and when other vehicles were moving from a driveway into the subject vehicle's path (15%).

Driver proficiency was less of an issue, accounting for only 17 percent of the incidents. Driver's aggressive driving behavior accounted for 3 percent of the incidents, and only 4 percent of incidents were drowsiness-related. Inattention to forward roadway was less of a factor in incidents than in near-crashes and crashes. Only 3 percent of incidents accounted for driving-related inattention. Distraction due to cell phone operations (6%) and passenger in vehicle (5%) were more representative than other distraction categories.

For avoidance maneuvers, a majority of drivers braked without locking up their wheels (58%), while others combined both braking and steering to left (15%) or braking and steering right (14%). Only one driver exhibited no reaction to another vehicle turning across the subject vehicle's path in the opposite direction.

The driving environment, as with crashes and near-crashes, was primarily clear weather (72%) during daylight hours (67%), with 70 percent of incidents in a business or industrial locality and 20 percent in residential areas. Out of 79, there was only 1 incident due to snow, and 7 percent were due to wet surface conditions, while the rest occurred on dry roads (91%). Approximately, 62 percent of incidents occurred at an intersection, and 11 percent were intersection-related. Over 92 percent of the incidents occurred on straight roads and 6 percent on curves. Only 4 percent of the incidents were due to sunlight glare, and there were no visual obstructions 91 percent of the time. Half of the incidents occurred in restricted traffic flow conditions, and 34 percent of the incidents occurred in free flow conditions. Four percent of incidents occurred due to roadway alignment.

Conflict with Vehicle in Adjacent Lane

Adjacent vehicle conflicts accounted for 1 percent of the crashes, 15 percent of the near-crashes, and 4 percent of the incidents. The conflict with an adjacent vehicle occurred more commonly when either the other vehicle was changing lanes ahead of the subject vehicle or when the subject vehicle was changing lanes.

Crash with Vehicle in Adjacent Lane. There was only one crash with a vehicle in an adjacent lane, and the incident type was left sideswipe in same direction with the subject vehicle changing lanes as the pre-incident maneuver.

The driver's aggressive behavior was the only factor in the left sideswipe collision with other vehicle in the adjacent lane. There were no other contributing factors. The driver did attempt to avoid the crash by braking and steering to the right.

For the associated factors, the collision occurred at an intersection on a straight, level business or commercial area in dry conditions during daylight hours and clear weather.

Near-Crashes with Vehicle in Adjacent Lane. Out of 115 near-crashes with an adjacent vehicle, 48 percent near-crashes occurred when other vehicles were changing lanes. When the other vehicle was changing lanes, 26 percent were with a right sideswipe threat and 14 percent were with a left sideswipe. There were 43 (37%) near-crashes when the subject vehicle was changing lanes with 19 percent being left sideswipe threat and 14 percent being a right sideswipe threat.

The subject vehicle or other vehicle changing lanes as the largest precipitating factor is not surprising considering blind spots and lane change maneuvers with more likelihood of left and right sideswipe threats with vehicles in another lane.

For driver contributing factors, 16 percent of near-crashes were classified as aggressive driving, and only 6 percent were classified as drowsiness-related. Nearly 42 percent of the near-crashes were classified as driver proficiency errors. Over one-fourth of the near-crashes had inattention to forward roadway as contributing factor. Four percent of the near-crashes were driving-related inattention, which included looking at the center mirror (1%), right window (2%) or left window (1%). The majority of secondary task inattention was due to a passenger in the vehicle (10%) and cell phone usage (7%), as shown in Figure 5.40.

Since the majority of near-crashes occurred when either another vehicle or the subject vehicle was changing lanes, the subject drivers mostly reacted by braking and steering either to left (31%) or right side (24%), while a few reacted by braking with no lockup (19%). Only 5 percent of drivers did not exhibit any avoidance maneuver.

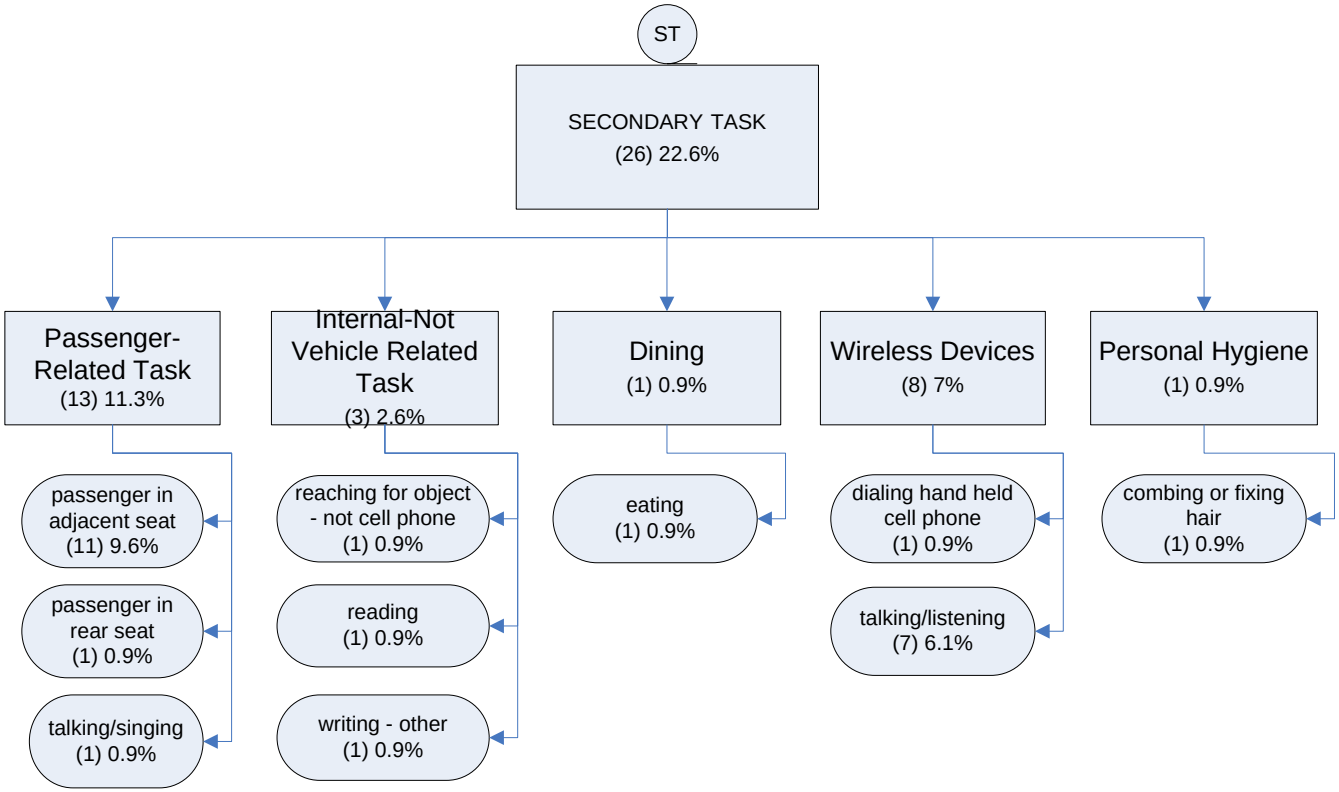


Figure 5.40. Breakdown of secondary tasks for near-crashes involving a vehicle in the adjacent lane.

The other contributing factors included sunlight glare (4%), rain (1%), a moving vehicle, (1%), roadway alignment (1%), and roadway delineation (2%).

When considering the driving environment associative factors, the road was wet/snowy in 13 percent of the near-crashes and it was raining in 8 percent of the near-crashes. Ninety percent of near-crashes were on straight roads and 10 percent were on curves. Nearly half of the near-crashes were in a business or commercial locality, while 40 percent were on interstate roadways. Thirty-one percent of near-crashes occurred in restricted traffic flow, while 28 percent occurred in free flow conditions. Moreover, 72 percent of near-crashes occurred in divided traffic flow conditions.

Incidents with Vehicle in Adjacent Lane. There were 342 incidents with other vehicles in an adjacent lane.

Similar to near-crashes, the biggest precipitating factors occurred when another vehicle was changing lanes (43%) and when the subject vehicle was changing lanes (34%). Only 8 percent of incidents corresponded to incidences when the subject vehicle moved off the roadway over the left lane line (3%) or over the right lane line (5%).

Driver proficiency was very similar between the near-crashes and incidents. Driver proficiency (33%), aggressive driving (16%) and driver drowsiness (6%) were all present for the incidents.

Inattention to forward roadway due to secondary tasks (15%) was less of a factor in incidents than in near-crashes with another vehicle in an adjacent lane. Distraction due to cell phone operations (4%) and a passenger in the vehicle (4%) were more representative than other distraction categories (Figure 5.41).

Eight percent of drivers had no avoidance maneuver, and 18 percent steered either to left or right without braking. Unlike with near-crashes, only a few drivers braked and steered to left (13%) or right (15%). The majority braked without locking up their wheels (33%). Interestingly only a few accelerated and steered either left (2%) or right (1%) as an avoidance maneuver.

However, there were more other contributing factors present in the incident than in the near-crashes. Roadway alignment (5%), road delineation (2%), weather-related visibility (1%), sunlight glare (4%), and traffic control devices (1%) were all present.

The driving environment, as with crashes and near-crashes, was primarily clear weather (82%) during daylight hours (77%), with 47 percent of incidents in a business or industrial locality and 38 percent on interstate roadways. Wet surface conditions (14%) were associated with some of the incidents. Compared to near-crashes, the intersection-related (7%) incidents were less of a factor. However, 11 percent of incidents occurred at entrance or exit ramps. Over 84 percent of the conflicts were on a straight road and 12 percent on a curve. Restricted traffic flow (41%) conditions were associated with the

incidents, and an additional 32 percent were in restricted speed conditions. Eighteen percent of incidents corresponded to conflicts at traffic signals, while others occurred when there were no traffic control devices (76%).

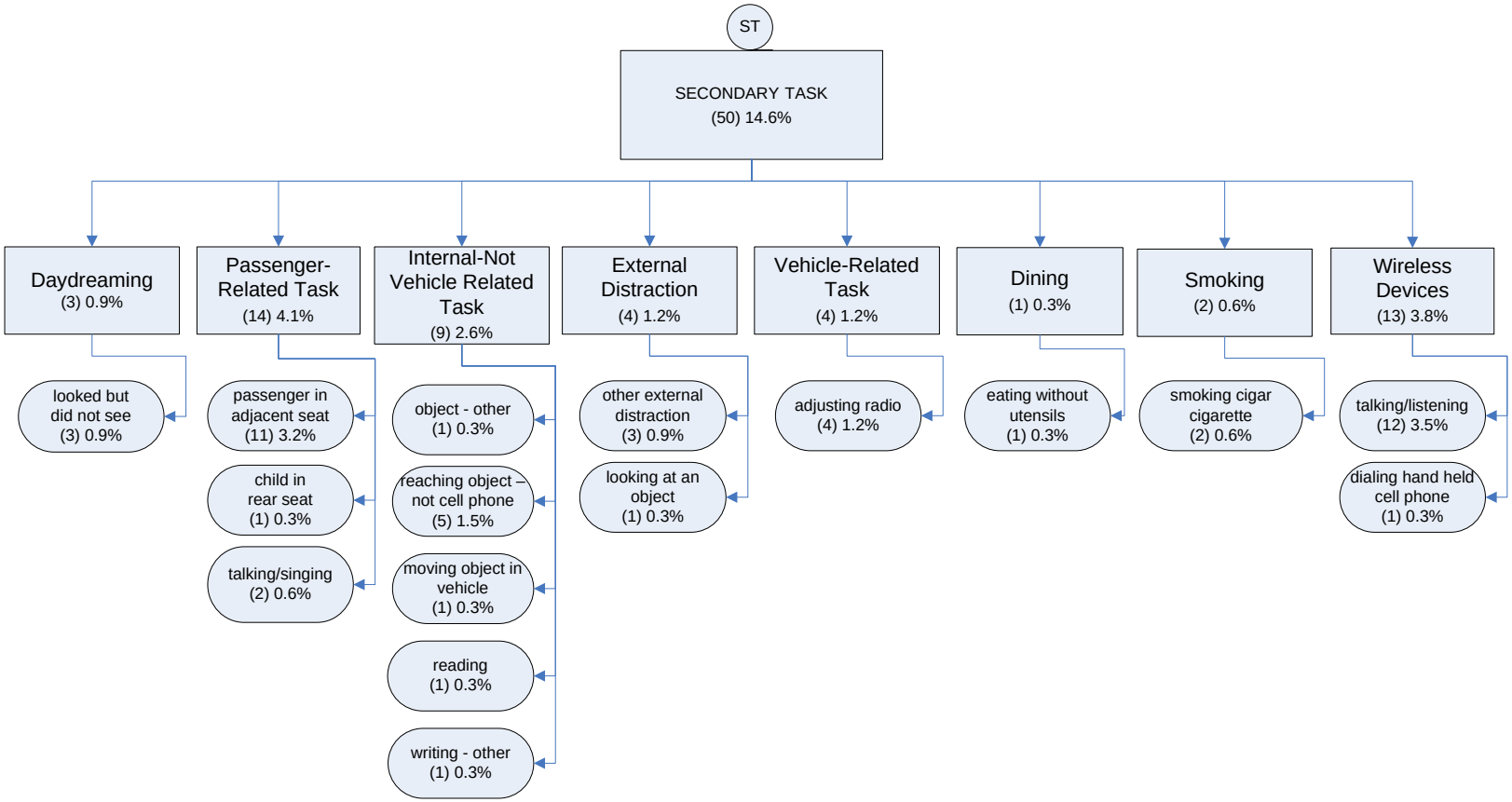


Figure 5.41. Breakdown of secondary tasks for incidents involving a conflict with vehicle in the adjacent lane.

Question 2. What Are The Relative Frequencies Of Primary And Contributing Factors For Each Level Of Severity?

All the contributing and associated factors used in this study were parsed into crashes and near-crashes. Tables are provided to group all the factors underneath the factor type categories listed in Table 5.7. This provides a mechanism to look at the relationship of these factors in crashes and near-crashes.

Table 5.7. Factor Type Categories.

Pre-event maneuver
Drivers' avoidance maneuver
Driver's willful behavior
Driver impairments
Driver proficiency error
Hands on wheel
Inattention to forward roadway
Surface condition
Relation to junction
Road alignment
Locality
Lighting
Visual obstruction
Weather
Trafficway flow
Traffic density
Traffic control device
Vehicle factors
Infrastructure

For the associative variables, GES data will be added to the tables to provide a comparison between the 100-Car Study data and GES data. This will only be done for the associative variables as these represent the entire event and are not dependent upon driver as are variables such as pre-event maneuver or avoidance maneuvers. While comparisons between avoidance maneuvers are possible between the GES and 100-Car Study database, this analysis would require more reduction and is beyond the scope of this current analysis. Table 5.8 below presents a comparison between the percentage of 100-Car Study crash and near-crash types to the percentage of GES crashes. These percentages demonstrate the generalizability of the 100-Car Study database to the GES crash database. Note that comparisons are presented for all 100-Car Study crashes, the 100-Car Study police reported crashes, and GES crashes. The results indicate that there is a large discrepancy between police-reported and non-police reported events.

Table 5.8. Comparison of crash types for 100-Car Crashes, Near-crashes, Police-Reported Crashes, and GES Crashes.

Conflict Type (100-Car)	100-Car Crash Frequency	100-Car Crash Percent	100-Car Police-Reported Crash Percent	100-Car Near-Crash Frequency	100-Car Near-Crash Percent	GES Crash Frequency	GES Crash Percent
Single Vehicle	24	34.8	8.3	48	6.3	2,740	4.6
Rear-End Striking and Rear-End Struck	27	39.1	66.7	450	59.1	14,722	24.9
Angle Collision	1	1.4	0.0	142	18.7	18,091	30.6
Sideswipe Opposite	2	2.9	16.7	27	3.5	480	0.8
Sideswipe Same	0	0.0	0.0	37	4.9	2,977	5.0
Head-On	0	0.0	0.0	27	3.5	1,879	3.2
Object/obstacle	9	13.0	8.3	6	0.8	11,063	18.7
Parked vehicle	4	5.8	0.0	5	0.7	2,027	3.4
Animal	2	2.9	0.0	10	1.3	1,515	2.6
Pedestrian	0	0.0	0.0	6	0.8	1,702	2.9
Pedalcyclist	0	0.0	0.0	0	0.0	1,085	1.8
Unknown	0	0.0	0.0	1	0.1	220	0.4
Other	0	0.0	0.0	2	0.3	655	1.1
Totals	**69**	**100%**	**100%**	**761**	**100%**	**59,156**	**100%**

Pre-Event Maneuver

When comparing crashes and near-crashes for the pre-event maneuver, the top three factors were ranked in the same order for both crashes and near-crashes (Table 5.9). These three factors accounted for 54 percent of the crashes and 80 percent of the near-crashes. For the top associated factor, 22 percent of crashes occurred when subject vehicle drivers were going straight at constant speed. The percentage of near-crashes (44%) was twice as high as that of crashes for this factor. The second highest factor, which was when the subject vehicle was decelerating in the traffic lane, was similar between crashes (22%) and near-crashes (20%). The third factor, going straight and accelerating, was slightly higher for near-crashes (16%) than crashes (10%). Turning right crashes (10%), the fourth highest factor, was more representative than turning right near-crashes (1%). When turning left, 6 percent were crashes, and 2 percent were near-crashes.

Table 5.9. Pre-event maneuver, crash, and near-crash.

Pre-event maneuver	Crash Frequency	Crash Percent	Near-Crash Frequency	Near-Crash Percent
Going straight, constant speed	15	21.7	331	43.5
Decelerating in traffic lane	15	21.7	155	20.4
Going straight, accelerating	7	10.1	120	15.8
Turning right	7	10.1	8	1.1
Stopped in traffic lane	5	7.3	24	3.2
Turning left	4	5.8	18	2.4
Starting in traffic lane	3	4.4	14	1.8
Merging	3	4.4	10	1.3
Changing lanes	2	2.9	60	7.9
Entering a parked position	2	2.9	0	0.0
Making U-turn	2	2.9	0	0.0
Maneuvering to avoid a vehicle	1	1.5	1	0.1
Backing up (not parking)	1	1.5	0	0.0
Leaving a parked position	1	1.5	0	0.0
Unknown	1	1.5	0	0.0
Negotiating a curve	0	0.0	20	2.6

Avoidance Maneuvers

Table 5.10 lists the different avoidance maneuvers used by the drivers. In 27 percent of the crashes there was no reaction; that is, the driver did not execute any avoidance maneuver. Only 4 percent of near-crashes had no reaction. Braking alone without lockup was the most common avoidance maneuver for near-crashes (45%) and crashes (25%). However, it was 20 percent more likely for near-crashes than crashes. Fifteen percent of drivers steered left as the avoidance maneuver in a crash. Only 6 percent of the near-crashes used this maneuver. The near-crashes (33%) were twice as likely to brake and steer as an avoidance maneuver as compared to the crashes (16%).

Table 5.10. Drivers' avoidance maneuver, crash, and near-crash.

No	Avoidance maneuver	Crash Frequency	Crash Percent	Near-crash Frequency	Near-crash Percent
1	No reaction	19	27.5	30	3.9
2	Braking(no lockup)	17	24.6	342	44.9
3	Steered to left	10	14.5	44	5.8
4	Braked and steered left	7	10.1	112	14.7
5	Braked and steered right	4	5.8	138	18.1
6	Braking(lockup unknown)	4	5.8	42	5.5
7	Unknown if action was attempted	3	4.4		
8	Steering to right	2	2.9	32	4.2
9	Other actions	2	2.9	5	0.7
10	Braking(lockup)	1	1.5	6	0.8
11	Accelerating and steering right			4	0.5
12	Accelerating			3	0.4
13	Accelerating and steering left			3	0.4

Willful Behavior

As shown in Table 5.11, aggressive driving (16%) contributed slightly more for crashes than for near-crashes (13%).

Table 5.11. Driver's willful behavior, crash, and near-crash.

No	Willful_Behavior	Crash Frequency	Crash Percent	Near-Crash Frequency	Near-Crash Percent
1	No willful behavior	58	84.1	658	86.5
2	Aggressive driving	11	15.9	103	13.4

Driver Impairments

As shown in Table 5.12, a total of 17 percent of crashes and 13 percent of near-crashes occurred with driver impairment as a contributing factor. Drowsiness was the most common impairment with 14 percent of the crashes and 11 percent of the near-crashes. Drugs or alcohol contributed to 2 percent of crashes.

Table 5.12. Driver impairments, crash, and near-crash.

No	Driver_Impairments	Crash Frequency	Crash Percent	Near-Crash Frequency	Near-Crash Percent
1	None apparent	23	39.0	400	54.4
2	Unknown	36	52.2	268	35.2
3	Drowsy, sleepy, asleep	8	13.6	79	10.7
4	Drugs, alcohol	1	1.7	0	0.0
5	Other	1	1.7	0	0.0
6	Angry	0	0.0	10	1.4
7	Other emotional state	0	0.0	4	0.5

Driver Proficiency

When comparing crashes and near-crashes, the driver proficiency error factor contributed to fewer crashes (28%) than near-crashes (43%). It is not clear why this 15 percent difference was present (Table 5.13).

Table 5.13. Driver proficiency error, crash, and near-crash.

No	Driver Proficiency	Crash Frequency	Crash Percent	Near-Crash Frequency	Near-Crash Percent
1	None	50	72.5	437	57.4
2	Driver proficiency error	19	27.5	324	42.6

Driver's Hands on Wheel

For the top associated factor, driving with left hand only, the percentages of crashes (30%) and near-crashes (32%) were similar (Table 5.14). When comparing crashes and near-crashes for hands-on-wheel categories, they were in the same ranked order. These three associated factors account for 41 percent of crashes and 53 percent of near-crashes. For the third factor, 25 percent of crashes occurred when subject vehicle drivers were driving with both hands. The percentage of near-crashes (35%) was higher than that of crashes for this factor. The percentage of near-crashes (16%) with right hand only was slightly higher than that of crashes (12%). The fifth factor, when subject vehicle drivers were driving with no hands on the wheel, was slightly more representative for crashes (5%) than near-crashes (3%).

Table 5.14. Hands on wheel, crash, and near-crash.

No	Hands on wheel	Crash Frequency	Crash Percent	Near-Crash Frequency	Near-Crash Percent
1	Left hand only	21	30.4	241	31.7
2	Unknown	20	29.0	115	15.1
3	Both hands	17	24.6	267	35.1
4	Right hand only	8	11.6	118	15.5
5	No hands on wheel	3	4.4	20	2.6

Driver Secondary Task Distraction and Inattention to Forward Roadway
For the distraction contributing factors shown in Table 5.15, more than one factor could be identified for each event. Therefore, the driver could be distracted both by talking on the cell phone and adjusting the radio during the event. For the purpose of this analysis, all the distractions were counted even if there were two for event. There were 2 distractions identified in 3 of the crashes and in 27 of the near misses. Distraction was not present in 39 percent of the crashes or in 61 percent of the near-crashes. This 22 percent difference may be why similar events became crashes instead of near-crashes.

Driving-related inattention to the forward roadway was a contributing factor in 14 percent of the crashes and 7 percent of the near-crashes. The largest portion of these crashes (10%) and near-crashes (3%) was attributable to looking out the left window.

The highest secondary task inattention to the forward roadway for crashes was drivers talking/listening cell phone (8%). Talking on a cell phone was the second highest distraction for near-crashes (4.8%). There were no crashes when drivers were either dialing a hand-held cell phone locating, reaching for, or answering a cell phone, operating a PDA, or performing other cell phone operations. On the other hand, over 3 percent of the near-crashes were due to these activities.

A passenger in the adjacent seat was the second highest crash secondary task contributing factor (7%) and the highest near-crash secondary task factor (6%). When passengers were seated in rear seats, drivers were involved in 1 percent of crashes and one near-crash. With a child in the rear seat there were no crashes and only one near-crash.

Animals or objects in the vehicle contributed to crashes (7%) as much as passengers in the adjacent seat. However, only 1 percent of the near-crashes were contributed to by this factor. An additional 3 percent of the crash inattention and 1 percent of the near-crash inattention was due to reaching for an object. For a more in depth discussion of distraction in crashes and near-crashes, read Chapter 7, *Goal 3* or Chapter 11, *Goal 7*.

Table 5.15. Inattention to forward roadway, crash, and near-crash.

	Crash Frequency	Crash Percent	Near-Crash Frequency	Near-Crash Percent
Not distracted	28	38.9	481	61.0
Left window	7	9.7	25	3.2
Talking/listening	6	8.3	38	4.8
Passenger in adjacent seat	5	6.9	48	6.1
No data	5	6.9	12	1.5
Animal/Object in Vehicle – Other	5	6.9	9	1.1
Reaching for object (not cell phone)	2	2.8	10	1.3
Cognitive – Other	2	2.8	5	0.6
Drinking from open container	2	2.8	1	0.1
Eating without utensils	1	1.4	15	1.9
Center mirror	1	1.4	14	1.8
Right window	1	1.4	14	1.8
Talking/singing	1	1.4	11	1.4
Other external distraction	1	1.4	10	1.3
Left mirror	1	1.4	9	1.1
Lost in thought	1	1.4	5	0.6
Moving object in vehicle	1	1.4	2	0.3
Passenger in rear seat	1	1.4	1	0.1
In-vehicle controls – Other	1	1.4	0	0.0
Dialing hand-held cell phone	0	0.0	14	1.8
Adjusting radio	0	0.0	10	1.3
Reading	0	0.0	10	1.3
Cell phone – Other	0	0.0	7	0.9
Other personal hygiene	0	0.0	7	0.9
Adjusting other devices integral to vehicle	0	0.0	5	0.6
Applying makeup	0	0.0	5	0.6
Dancing	0	0.0	3	0.4
Locating/reaching/answering cell phone	0	0.0	2	0.3
Looked but did not see	0	0.0	2	0.3
Right mirror	0	0.0	2	0.3
Child in rear seat	0	0.0	1	0.1
Combing or fixing hair	0	0.0	1	0.1
Drinking	0	0.0	1	0.1
Eating	0	0.0	1	0.1
Insect in vehicle	0	0.0	1	0.1
Inserting/retrieving CD	0	0.0	1	0.1
Looking at an object	0	0.0	1	0.1
Looking at pedestrian	0	0.0	1	0.1
Looking at previous crash or incident	0	0.0	1	0.1
Operating PDA	0	0.0	1	0.1
Smoking cigar/cigarette	0	0.0	1	0.1

Visual Obstructions
The visual obstructions category included factors that contributed in some way to the crashes or near-crashes. Not surprisingly most of the crashes (86%) and near-crashes (89%) had no visual obstruction. Only 7 total crashes had visual obstruction contributing factors. Reflected glare contributed to two crashes but did not seem to contribute near-crashes. On the other hand 5 percent of the near-crashes had sunlight as a contributing factor, but only one of the near-crashes did. Moving and parked vehicle visual obstructions were similarly representative between crashes and near-crashes (Table 5.16).

Table 5.16. Visual obstruction, crash, and near-crash.

No	Visual obstructions	Crash Frequency	Crash Percent	Near-Crash Frequency	Near-Crash Percent
1	No obstruction	59	85.5	675	88.7
2	Reflected glare	2	2.9	2	0.3
3	Sunlight glare	1	1.4	39	5.1
4	Moving vehicle	1	1.4	18	2.4
5	Parked vehicle	1	1.4	9	1.2
6	Other obstruction	1	1.4	3	0.4
7	Trees, crops, vegetation	1	1.4	0	0.0
8	Rain, snow, fog, smoke, sand, dust	0	0.0	4	0.5
9	Roadway infrastructure such as building, billboard, signs, embankments, etc.	0	0.0	3	0.4
10	Curve/hill	0	0.0	2	0.3
11	Headlight glare	0	0.0	1	0.1
12	Unknown	3	4.3	5	0.7

Road Surface Condition
When comparing crashes and near-crashes on different surface conditions, the top four factors have the same ranked order for both crashes and near-crashes in the 100-Car Study database and GES database. Not surprisingly, dry surface conditions were the most common associated factor, with fewer crashes in both the 100-Car Study database (74%) and GES database (76%) than near-crashes (86%). In the wet, snowy, and icy conditions, there were more crashes, both 100-Car Study and GES, in each factor than near-crashes. Even though these were not necessarily classified as contributing factors, it seems that the reduced traction from non-dry roads may have contributed to these events becoming crashes instead of remaining near-crashes (Table 5.17).

Table 5.17. Surface condition, crash, and near-crash.

No	Surface condition	Crash Frequency	Crash Percent	Near-Crash Frequency	Near-Crash Percent	GES Crash Frequency	GES Crash Percent
1	Dry	51	73.9	654	85.9	45171	76.4
2	Wet	13	18.8	98	12.9	10039	17.0
3	Snowy	4	5.8	4	0.5	1766	3.0
4	Icy	1	1.5	4	0.5	1347	2.3
5	Unknown	0	0.0	1	0.1	692	1.2

Relation to Junction

When comparing crashes and near-crashes for relation to junction, note that the general relationship between all variables is the same for 100 Crashes, GES crashes, and 100-Car Study near-crashes is similar. The top 5 factors have the same ranked order for crashes in both 100-Car Study and GES database and 100-Car Study near-crashes (Table 5.18). The 5 associated factors account for 96 percent of 100-Car Study crashes, 93 percent of GES crashes, and 97 percent of near-crashes. For the biggest factor, non-junction, the percentage of near-crashes (60%) was one-half times higher than 100-Car Study crashes (38%) and slightly less than one-half than the GES crashes (47%). Intersection crashes in both 100-Car Study (25%) and GES (23.5%), the second highest factor, was more representative than near-crashes (20%). The intersection-related events, the third highest factor accounted for 100-Car Study crashes (16%) and GES crashes (15%), was one and one-half times higher than near-crashes (10%). For the fourth highest factor, the entrance or exit ramp 100-Car Study crashes (9%) and GES crashes (0.2%) did not match but 100-Car Study crashes were more similar to near-crashes (5%). In parking lots, the percentage of 100-Car Study crashes (9%) and GES crashes (8%) was four times higher than that of near-crashes (2%). Note that GES had a relatively high percentage of crashes on interchanges that the 100-Car Study database did not. This could be due, in part, to difficulty in determining whether the vehicle was in an "interchange" using video only.

Table 5.18. Relation to junction, crash, and near-crash.

No	Relation to Junction	Crash Frequency	Crash Percent	Near-Crash Frequency	Near-Crash Percent	GES Crash Frequency	GES Crash Percent
1	Non-junction	26	37.7	456	59.9	27498	46.5
2	Intersection	17	24.6	149	19.6	13904	23.5
3	Intersection-related	11	15.9	76	10	8989	15.2
4	Entrance/exit ramp	6	8.7	40	5.3	133	0.2
5	Parking lot	6	8.7	14	1.8	4437	7.5
6	Driveway, alley access, etc.	2	2.9	8	1.1	0	0.0
7	Other	1	1.5	1	0.1	759	1.3
8	Interchange area	0	0.0	16	2.1	2907	4.9
9	Unknown	0	0.0	1	0.1	529	0.9

Roadway Alignment

As expected, straight, level roads and curved, level roads accounted for the majority of 100-Car Study crashes (94%) and near-crashes (97%) but only 49 percent of GES crashes. This is due primarily to lack of GES information as 35 percent of all GES crashes, alignment is unknown. Further comparisons of GES to 100-Car Study data will not be made for the roadway alignment category. The straight, level roads factor for near-crashes (84%) was slightly higher than for 100-Car Study crashes (75%), whereas the curve level for 100-Car Study crashes (19%) was slightly higher than for near-crashes (13%). The curve grade was also higher for crashes (4%) than for near-crashes (one%). Although not a strong association, it is interesting that curves were more associated with crashes than with near-crashes (Table 5.19).

Table 5.19. Road alignment/road profile, crash, and near-crash.

No.	Road alignment/road profile	Crash Frequency	Crash Percent	Near-crash Frequency	Near-Crash Percent	GES Crash Frequency	GES Crash Percent
1	Straight level	52	75.4	638	83.8	26265	44.4
2	Curve level	13	18.8	99	13	2898	4.9
3	Curve grade	3	4.4	7	0.9	3048	5.2
4	Straight grade	1	1.5	15	2	5582	9.4
5	Unknown	0	0.0	1	0.1	20460	34.6
6	Straight hillcrest	0	0.0	1	0.1	594	1.0

Locality of Event

The "locality of event" variable was adopted from the Virginia State Police Accident Report, not the GES database; therefore, no GES data will be presented for this variable. When comparing crashes and near-crashes in different localities, the top two factors are very similar. These two factors account for 64 percent of crashes and 62 percent of near-crashes. The business or industrial area was the most common location for both crashes (45%) and near-crashes (44%). The business or industrial area is likely a common driving environment in the northern Virginia area, and these large percentages are not surprising. The second highest factor, driving on open country roads, was the same between crashes (18%) and near-crashes (18%). The percentage of crashes (16%) in residential areas was twice the percentage of near-crashes (8%). On the other hand, the near-crashes (28%) that occurred on interstate roads were more representative than crashes (12%). This difference between crashes and near-crashes is likely due to the higher percentage of rear-end near-crashes as compared to crashes (Table 5.20). These rear-end near-crashes are likely to be more associated with interstate driving.

Table 5.20. Locality, crash, and near-crash.

No	Locality	Crash Frequency	Crash Percent	Near-Crash Frequency	Near-Crash Percent
1	Business/industrial	31	44.9	335	44.0
2	Open Country	13	18.8	138	18.1
3	Residential	11	15.9	60	7.9
4	Interstate	8	11.6	212	27.9
5	Other	5	7.3	3	0.4
6	Construction zone	1	1.5	11	1.5
7	Church	0	0.0	1	0.1
8	Unknown	0	0.0	1	0.1

Lighting

As shown in Table 5.21, the top two lighting factors are ranked in the same order for 100-Car Study crashes, GES crashes, and 100-Car Study near-crashes. These two factors account for 87 percent of 100-Car Study crashes, 84 percent of GES crashes, and 83 percent of near-crashes. Most 100-Car Study crashes (62%), GES crashes (67%) and near-crashes (66%) occurred during day or daylight conditions. Driving in darkness with lighted conditions was more representative for 100-Car Study crashes (25%) than GES crashes (17%) or near-crashes (17%). The unlit

darkness factor was similar between crashes (7%) and near-crashes (7%) but was almost twice as high for GES crashes (12%). For the dusk factor the percentage of near-crashes (9%) was twice as high as the percentage of crashes (4%) and even more for GES crashes (2%).

Table 5.21. Lighting, crash, and near-crash.

No	Lighting	Crash Frequency	Crash Percent	Near-crash Frequency	Near-crash Percent	GES Crash Frequency	GES Crash Percent
1	Daylight	43	62.3	502	66	39526	66.8
2	Darkness lighted	17	24.6	126	16.6	9930	16.8
3	Darkness not lighted	5	7.3	54	7.1	7040	11.9
4	Dusk	3	4.4	65	8.5	1302	2.2
5	Dawn	1	1.5	14	1.8	932	1.6

Weather

In the weather category, the clear weather factor was associated with the same percentage of 100-Car Study crashes (78%), GES crashes (83%), and near-crashes (78%). The second most associated weather factor was rain. When raining, the drivers had a slightly higher percentage of crashes (12%) than near-crashes (8%). The 100-Car Study crash percentage (12%) and GES crash percentage (11%) are similar. When comparing the 100-Car Study crashes and GES crashes to near-crashes, the differences may be due to the potentially reduced traction and visibility associated rain playing a role in drivers' inability to avoid a crash. Snow followed a similar pattern, however only one 100-Car Study crash occurred with snow as an associated factor. Cloudy weather was associated with more near-crashes (13%) than crashes (9%) (Table 5.22).

Table 5.22. Weather, crash, and near-crash.

No	Weather	Crash Frequency	Crash Percent	Near-crash Frequency	Near-crash Percent	GES Crash Frequency	GES Crash Percent
1	Clear	54	78.3	599	78.7	49107	83.0
2	Raining	8	11.6	57	7.5	6616	11.2
3	Cloudy	6	8.7	99	13	0	0.0
4	Snowing	1	1.5	3	0.4	1915	3.2
5	Fog	0	0.0	1	0.1	218	0.4
6	Mist	0	0.0	1	0.1	0	0.0
7	Unknown	0	0.0	1	0.1	835	1.4

Trafficway Flow

Not surprisingly, the majority of crashes in the 100-Car Study, GES crashes, and near-crashes occurred on divided or non-divided trafficways because these are the most common types of roadways. As shown in Table 15, the non-divided trafficway was associated with more 100-Car Study crashes (46%) and GES crashes (49%) than the near-crashes (36%). On the other hand more near-crashes (59%) were associated with divided roadways than 100-Car Study crashes (42%) or GES crashes (32%). The third highest associated factor (one-way traffic flow)

accounted for 9 percent of 100-Car Study crashes, only 5 percent of GES crashes, and 3 percent of the near-crashes. For the fourth factor, when the subject vehicle drivers were driving in traffic with no lanes, the percentage of crashes (three%) was three times higher than the percentage of near-crashes (one%). Please note that 15 percent of all GES crashes, the trafficway flow variable is unknown (Figure 5.23).

Table 5.23. Trafficway flow, crash, and near-crash.

No	Trafficway_Flow	Crash Frequency	Crash Percent	Near-crash Frequency	Near-crash Percent	GES Crash Frequency	GES Crash Percent
1	Not divided	32	46.4	277	36.4	29,001	49.0
2	Divided (median strip or barrier)	29	42	449	59	18,609	31.5
3	One-way traffic	6	8.7	26	3.4	2,925	4.9
4	No lanes	2	2.9	8	1.1	0	0.0
5	Unknown	0	0.0	1	0.1	8,621	14.6

Traffic Density

The traffic density variable (Level of Service variable) is also not used in the GES database; therefore, no GES comparisons will be discussed. When comparing crashes and near-crashes for the traffic density, all seven factors are the same rank order for both crashes and near-crashes (Table 5.24). All these factors account for 100 percent of crashes and 100 percent of near-crashes. Free flow was the top associated factor, with crashes (60%) being almost twice as represented as near-crashes (32%). This difference between crashes and near-crashes is likely due to the higher percentage of rear-end near-crashes as compared to crashes. These rear-end near-crashes are likely to be more associated with flow restrictions. The higher percentage of near-crashes as compared to crashes is shown for the first four flow restrictions. It is also interesting that these first 5 flow restrictions are also ranked in Table 5.24 with each consecutive factor being more restrictive.

Table 5.24. Traffic density, crash, and near-crash.

No	Traffic density	Crash Frequency	Crash Percent	Near-Crash Frequency	Near-Crash Percent
1	Free flow	41	59.4	244	32.1
2	Flow with some restrictions	14	20.3	233	30.6
3	Stable flow, maneuverability and speed more restricted	7	10.1	191	25.1
4	Unstable flow, temporary restrictions slow driver	4	5.8	64	8.4
5	Flow is unstable, vehicles are unable to pass temporary stoppages, etc.	2	2.9	26	3.4
6	Forced traffic flow condition with low speeds and traffic volumes that are below capacity	1	1.5	2	0.3
7	Unknown	0	0.0	1	0.1

Traffic Control Devices
Almost twice as many crashes (46%) in the 100-Car Study had traffic control devices than did near-crashes (24%) (Table 5.25). GES crashes were not similar to either 100-Car Study crashes or near-crashes as 33 percent of all GES crashes occurred in the presence of a traffic control device. Traffic signals were the most common traffic control device for 100-Car Study crashes (28%), GES crashes (20%), and near-crashes (17%). The remaining traffic control devices were similar between 100-Car Study crashes than near-crashes. The stop signs accounted for three times more 100-Car Study crashes (6%) and 5 times more for GES crashes (10%) than the near-crashes (2%). When comparing the crashes and near-crashes for traffic control device, the fourth, fifth, and sixth factors in Table 5.24 are the same rank order for 100-Car Study crashes, GES crashes, and near-crashes. These three factors account for 10 percent of the crashes 3 percent of GES crashes and 5 percent of the near-crashes. These three factors were: traffic lanes marked; other; and yield signs. Note that for 4 percent of GES crashes, the presence of a traffic control device is unknown.

Table 5.25. Traffic control device, crash and near-crash.

No	Traffic Control Device	Crash Frequency	Crash Percent	Near-crash Frequency	Near-crash Percent	GES Crash Frequency	GES Crash Percent
1	None	37	53.6	574	75.4	36,506	61.7
2	Traffic signal	19	27.5	130	17.1	11,709	19.8
3	Stop sign	4	5.8	16	2.1	5,954	10.1
4	Traffic lanes marked	3	4.4	20	2.6	0	0.0
5	Other	2	2.9	9	1.2	1,179	2.0
6	Yield sign	2	2.9	8	1.1	637	1.1
7	Officer or watchman	2	2.9	1	0.1	119	0.2
8	Unknown	0	0.0	2	0.3	2,456	4.2
9	Slow or warning sign	0	0.0	1	0.1	596	1.0

Vehicle Factors
For vehicle factors, some unknown factor contributed to 2 percent of crashes and even smaller percentage of near-crashes (0.1%). Due to the small numbers in the 100-Car Study database, no comparisons will be made between the two databases (Table 5.26).

Table 5.26. Vehicle factors, crash and near-crash.

No	Vehicle_factors	Crash Frequency	Crash Percent	Near-crash Frequency	Near-crash Percent
1	None	68	98.6	760	99.9
2	Unknown	1	1.5	1	0.1

Infrastructure
This infrastructure variable was adopted from the Light Vehicle/Heavy Vehicle technical report (Hanowski, et al, 2000). GES does not use this variable; therefore no comparisons can be made. With infrastructure, the majority of crashes (80%) and the majority of near-crashes had no

infrastructure contributing factor (Table 5.27). When comparing the crashes and near-crashes, the top two contributing factors for each were in the same rank order. These two factors (roadway geometry and delineation) accounted for 15 percent of crashes and 3 percent of near-crashes. The roadway sight distance factor was not representative of crashes, and it contributed to less than 1 percent of near-crashes.

Table 5.27. Infrastructure, crash and near-crash.

No	Infrastructure	Crash Frequency	Crash Percent	Near-Crash Frequency	Near-Crash Percent
1	None	55	79.7	727	95.5
2	Roadway geometry	5	7.3	17	2.2
3	Roadway delineation	5	7.3	7	0.9
4	Traffic control device	1	1.5	1	0.1
5	Roadway sight distance	0	0.0	4	0.5
6	Unknown	3	4.4	5	0.7

Question 3. What are the dynamic reconstructions of each crash and near-crash, and what are the stimulus-response times associated with each?

In addition to the analysis of coded data relevant to near-crashes and crashes, dynamic reconstructions and animated representations of the events accompany this document. The reconstructions span the period from 10 seconds prior to impact (or successful crash avoidance) to 2 seconds after impact or avoidance. A scenario timeline is provided that captures vehicle pre-event actions, driver actions, speeds, ranges, range rates, braking, steering (if applicable), and trajectories.

By way of a brief description for the crashes, Table 5.28 is provided. This table includes event nature, crash number, a narrative description, and a graphical depiction of each crash. These 33 crashes include all level I, II, and III crashes for which video was available (level IV crashes were not included).

This page intentionally left blank.

CHAPTER 6: GOAL 2, OPERATIONALLY DEFINE A NEAR-CRASH USING QUANTITATIVE MEASURES

DATA ANALYSIS OVERVIEW

For this study, near-crashes and crashes were operationally defined based upon the *a priori* criteria described below:

- Crash: Any contact with an object, either moving or fixed, at any speed in which kinetic energy is measurably transferred or dissipated. Includes other vehicles, roadside barriers (curbs and tire strikes), objects on or off of the roadway, pedestrians, cyclists, or animals.
- Near-Crash: Any circumstance that requires a rapid, evasive maneuver by the subject vehicle, or by any other vehicle, pedestrian, cyclist, or animal, to avoid a crash. A rapid, evasive maneuver is defined as steering, braking, accelerating, or any combination of control inputs that approaches the limits of the vehicle capabilities. As a guide, a subject vehicle braking greater than 0.5 g or steering input that results in a lateral acceleration greater than 0.4 g to avoid a crash, constitutes a rapid maneuver.

As shown, while these criteria were based somewhat upon quantitative kinematic criteria, they were subjective in nature. While such definitions were useful for purposes such as classifying video data, they were not useful for precisely defining events or as criteria for other purposes, such as warning algorithms. Therefore, a goal of the 100-Car Study, given that it contains more crash, near-crash, and incident data than ever before collected, was to explore the feasibility of creating more useful operational definitions of near-crash events.

Near-crashes can be defined quantitatively based upon time-to-collision, acceleration, or proximity criteria. However, the results of this and other studies have shown that there is inherent *"noise"* present in such criteria. Thus, there are inherent difficulties associated with both quantitative and qualitative approaches. For example, the qualitative definitions required explanations in quantitative terms (How rapid is a *"rapid maneuver"?),* and the quantitative definitions at some level must be based upon a subjective assessment (Does a TTC of 0.1 s second constitute a near-crash when such values occur regularly on interstates at rush hour?).

Indeed, all attempts to quantify events strictly on the basis of somewhat simple quantitative kinematic criteria for this study led to a number of false positives. As described in more detail as part of Chapter 13, *Goal 9*, there were many instances in which a sensor provided a data signature that was misinterpreted by an algorithm. In addition, there were many instances in which a *"normal"* driving maneuver by a participant produced a kinematic signature that was virtually identical to the criteria used to identify a near-crash.

The most common example of a false sensor signature was a misinterpreted radar target. Despite considerable time and effort spent attempting to filter radar data, there were many occasions in which the identified target was not in the vehicle travel lane due to road geometry or other factors. While such events could be readily identified as being "false positives" upon video review of the event, filtering based upon objective kinematic data alone was more problematic. This finding and some of the methods used to help alleviate the problem are discussed in more detail below.

The most common example of a driver behavior that mimicked a near-crash signature was a *"flying pass."* A typical *"flying pass"* event occurred when a driver rapidly approached a string of stopped vehicles and made a planned lane change into a right/left dedicated turning lane. From the perspective of reviewing near-crash signatures, this scenario produced a very high range rate and a very short range, as would a near-crash event. However, upon video review it could be readily seen that the driver is fully alert, making a planned maneuver and taking very little risk, particularly when the lead vehicle had completely stopped and the range could be accurately gauged.

Consequently, the experience of this study was that qualitative and quantitative criteria were dependent upon one another to some degree. A qualitative criterion must incorporate both quantitative criteria as well as crash risk. Similarly, a quantitative criterion alone will not suffice without qualitative information regarding the validity of the near-crash based upon context information such as the presence of a planned versus an unplanned maneuver. Use of both quantitative and qualitative criteria led to the creation of a database of *"valid"* and *"invalid"* (i.e., false positive for reasons of sensor or behavior) events that could be used to test and refine classification criteria to quantitatively define and capture near-crashes.

Several recent studies (e.g., Smith, Najm, and Lam, 2003) have attempted to quantify safety surrogates including near-crash and less severe events using quantitative criteria, such as range and range rate to create near-crash and conflict boundaries. Some of these approaches used range/range rate trajectories, while others used a single-point approach that represented the greatest crash threat in a trajectory. For the purposes of this analysis, we did not distinguish between the two approaches mathematically since the greatest threat in a range/range rate trajectory would drive the categorization of an event. We did, however, include graphs depicting both approaches for two reasons: (1) more sophisticated analysis can be employed using the trajectories (e.g., by requiring some degree of sample continuity to filter radar data), and, as a result, it may be useful to see the trace data; and (2) trace data are sometimes hard to follow graphically and the point data often aids in visualization.

Data Included in the Analyses

The data included in this section and used for the analyses to define near-crashes represents only conflicts with lead and following vehicles and not the other conflict types. There were two reasons for this: radar data for all 100 cars was only available for the front and rear of the vehicle, and it proved to be very difficult to discriminate between valid and invalid events for dynamically complex events. A dynamically complex near-crash has been operationally defined for this analysis as an event in which the kinematic data is incomplete or unclear due to sensor availability or signal quality. An example of a dynamically complex near-crash would be an event for which the driver is making a lane change in stop-and-go traffic. A dynamically simple near-crash is one for which the kinematic data is complete and clear (e.g., a rear-end striking conflict for which the range-range/rate information and video is present for the duration of the conflict). In other words, what was defined as simple or complex was an artifact of the vehicle sensor suite and system capabilities.

The range/range rate quantification approach analyzed lead and following-vehicle conflicts separately. We assessed the threat points and trajectories for the valid crashes and near-crashes as well as all of the invalid events. The data were constrained in the following manner:

1) The most relevant (closest range) target was used to compute time-to-collision using a constant acceleration equation.
2) The lead or following vehicle had a range that was less than 50 meters at some point in the trajectory.
3) The range rate became negative for at least one frame (1/10 second).

For the lead-vehicle conflict case, the resulting dataset contained 11 valid crash events and 290 valid near-crash events for a total of 309 valid events. Also used in the lead-vehicle analysis were 6,186 invalid events. For the following-vehicle conflict case, the data contained 9 crash events and 52 near-crash events for a total of 61 valid events. There were 157 invalid events used in the following-vehicle analysis.

Question 1. What kinematic variables best predict the occurrence of crashes and near-crashes?

Conflicts with Lead vehicles

Using the range/range rate approach, the valid and invalid events were categorized based upon a number of criteria. First, crash and near-crash boundaries were estimated in terms of range/range-rate based on a forward collision warning project conducted by the Collision Avoidance Metrics Partnership (CAMP) (Kiefer et al., 2003). Since the boundary equations were not provided in the cited paper, these boundaries were estimated from graphs shown in the paper. The approximations were:

Warning Boundary: Range = -RangeRate * 3.5
Conflict Boundary: Range = -RangeRate * 4 + 10

It is important to note that while these were good approximations, they do not represent the exact curves provided in the paper. This is particularly true of the conflict boundary equation, which appeared to have a second order term.

The purpose of the boundaries described above was to provide criteria for forward crash warnings for drivers. As such, they necessarily must weight the cost of a *"miss"* (i.e., no warning provided when a threat is present) much higher than a false alarm (i.e., a warning is provided in which there is no imminent threat).

Another purpose of creating quantitative criteria for a near-crash was for the potential detection of near-crash events in large naturalistic driving databases. In this case, the detection criteria did not need to be weighted so heavily toward very few misses. However, missing near-crash or crash events is also not desirable in naturalistic data collection, even though it does not have the safety implications present in forward crash warning systems. Thus, a third boundary was calculated that attempted to minimize the overall error rate (i.e., misses + false alarms) to show what might be possible for this application. The equation used for this boundary condition for the lead-vehicle conflict case was:

Minimum Error Boundary: Range = -RangeRate*1.3x + 0.5

Note that more sophisticated multivariate modeling to discriminate between valid crash and near-crashes as well as invalid events for the purposes of triggering event data collection for a large-scale study is included as part of Chapter 13, *Goal 9,* Determine Rear-End Contributing Factors and Dynamic Conditions.

A separate analysis was also conducted as part of this goal to determine the degree to which the three criteria described above could capture the crash events collected in the 100-Car Study. For this analysis, the crash trajectories were stopped 2.0 seconds prior to the crash event to determine whether or not the trajectory had crossed each of the boundaries. This provided some insight into the validity of the various approaches for detecting crash events in sufficient time to warn a driver.

The analysis of the lead-vehicle conflict case using the approaches described above is shown in Table 6.1. The data used in this analysis included: 11 valid crash events, 290 valid near-crash events, and 6,186 invalid events. It is important to understand the nature of an invalid event in this context. An invalid event is an event that was triggered by a signature associated with a possible lead-vehicle event. These event triggers included short time-to-collisions, high longitudinal decelerations, or some combination of the two. As described above in Chapter 2, trained analysts reviewed each of these events and determined that a conflict was not present. However, these events should not be construed as "normal" driving cases and instead represented the most difficult cases for discrimination purposes since they themselves represented the extremes of range and range rate from the roughly 43,000 hours of data collected for this study.

Table 6.1. Percentage of hits and false alarms for each boundary model for a conflict with a lead vehicle.

Conflict with Forward Vehicle	Valid crash + near-crash hit rate	Invalid false alarm rate	Diff	Crash hit rate
Minimum Error	74%	20%	54%	10 / 11
~Warning	90%	73%	17%	10 / 11
~Conflict	97%	97%	0%	11 / 11

Some of the lead- and following-vehicle crashes and near-crashes used in the calculations shown in Table 6.1 were discovered using triggers other than those based upon time-to-collision. This was because in some cases the radar did not correctly identify the crash/near-crash target that constituted the greatest threat. Thus, some of the events were captured by deceleration (or other) triggers. In these cases, the points may have been misclassified based on range/range rate calculations and therefore were not accurate depictions of the actual threat. Although the exact number of these points was not currently known, it was estimated to be in the range of 10 percent.

As shown in Table 6.1, the approximations of the CAMP warning and conflict boundaries provided very high hit rates for both crash and near-crash events and detected all of the crashes with the exception of one case. However, the false alarm rate was also very high, which indicated difficulty in discriminating valid versus invalid events as defined here. In contrast, the

minimum error boundary had a much lower false alarm rate, but at the expense of a 26 percent miss rate.

The results in Table 6.1 are depicted graphically in Figure 6.1. Each point on the graph represents the point of greatest threat with the lead vehicle, which is defined as the moment at which the two vehicles are closest during the event as determined by the radar signature. The red points represent a random sample of the invalid events and the blue points are valid crash and near-crash events. The area above each boundary line indicates events that would not be triggered using the kinematic equation represented by each boundary. Alternatively, the area below each boundary line indicates events that would be triggered. The red points above a boundary represent correct rejections. The red points below the line are false alarms. Likewise, the blue points above a boundary line represent misses, while blue points below a line would indicate a hit (correctly providing a warning).

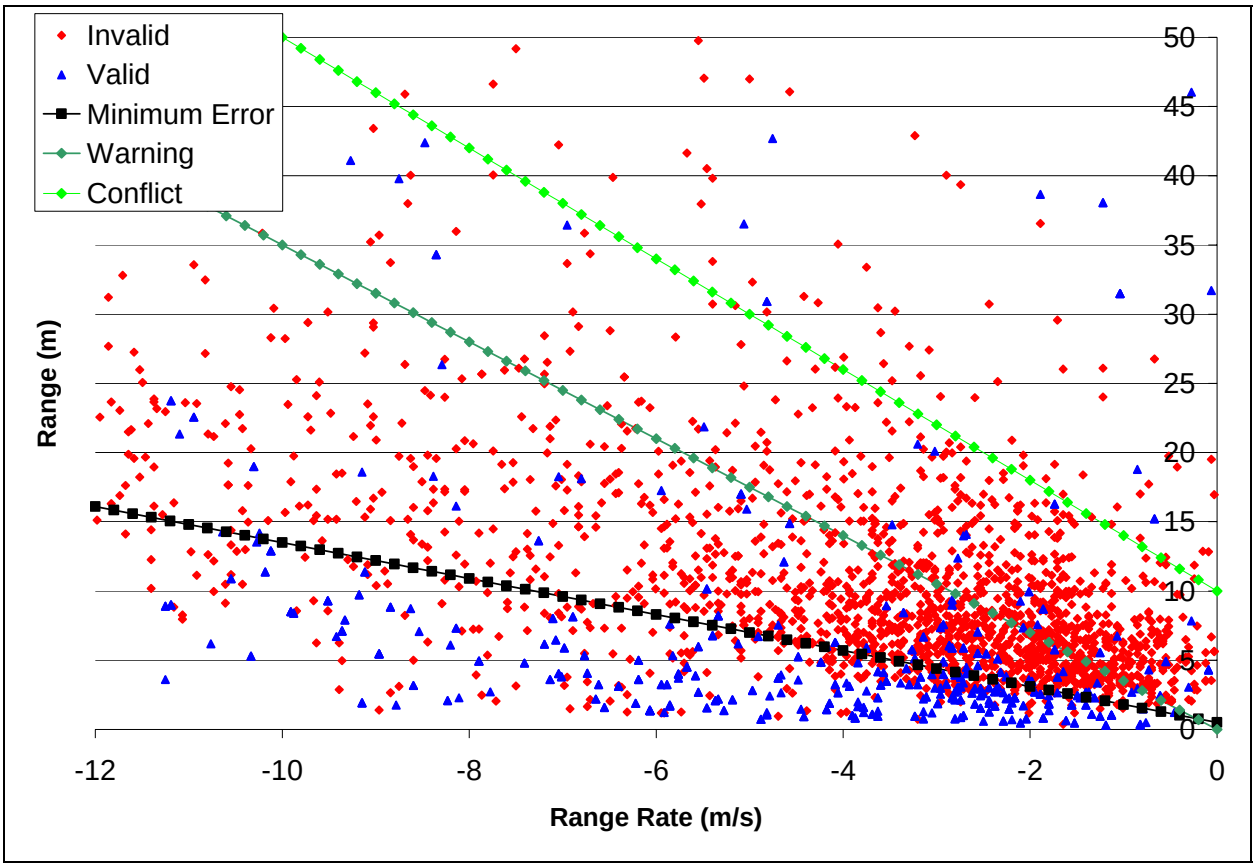


Figure 6.1. Point of greatest threat with lead vehicle for all crashes and near-crashes, and a random sample of invalid events. The boundaries shown are approximations of the warning and conflict boundaries used as part of the forward collision warning algorithm (Kiefer et al., 2003) and a minimum error boundary calculated for this dataset.

As another means to visualize the data, Figure 6.2 shows the trajectories of the selected events. Trajectories show the timeline of a vehicle for up to 8 seconds prior to the trigger. The point at which the trajectory crosses a boundary is the point of a hypothetical warning. As before, the

color blue indicates a valid event and the color red indicates an invalid event. A blue line crossing the boundary would be considered a correctly identified valid event. A red trajectory crossing a boundary would be a false alarm. Blue trajectories that do not cross a boundary are misses. To enhance data visualization, Figure 6.3 provides a random sample of the trajectories.

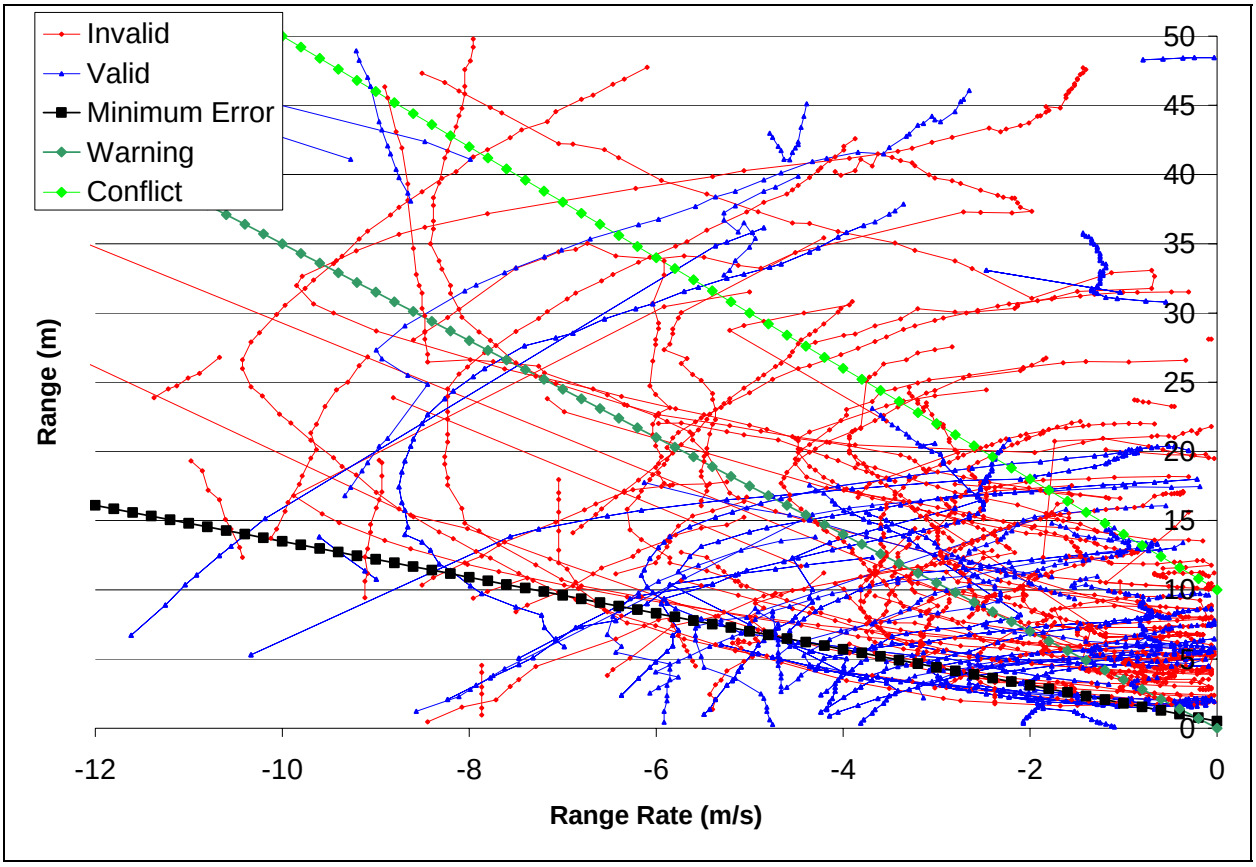


Figure 6.2. Range/range rate trajectories of vehicles approaching a lead vehicle including crashes and near-crashes, and a random sample of invalid events. The boundaries shown are approximations of the warning and conflict boundaries used as part of the forward collision warning algorithm (Kiefer et al., 2003) and a minimum error boundary calculated for this dataset.

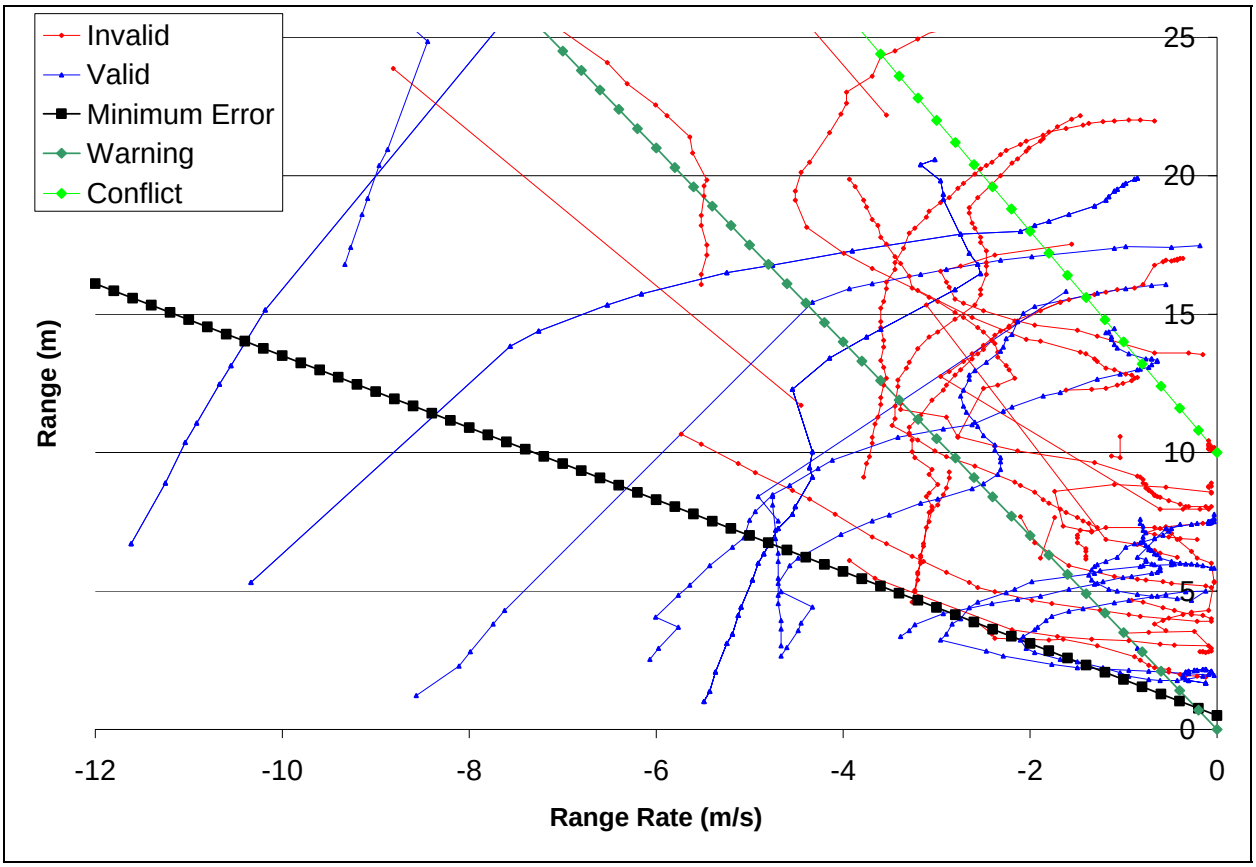


Figure 6.3. Range/range rate trajectories of vehicles approaching a lead vehicle. This data includes a random sample of crashes and near-crashes, and invalid events to improve visualization. The boundaries shown are approximations of the warning and conflict boundaries used as part of the forward collision warning algorithm (Kiefer et al., 2003) and a minimum error boundary calculated for this dataset.

With regard to warning timing, Figure 6.4 provides range/range rate trajectories for the crash events. The large blue points on each trajectory represent the time during the trajectory that was at least 2 seconds prior to the collision and during which the crash could have been predicted.

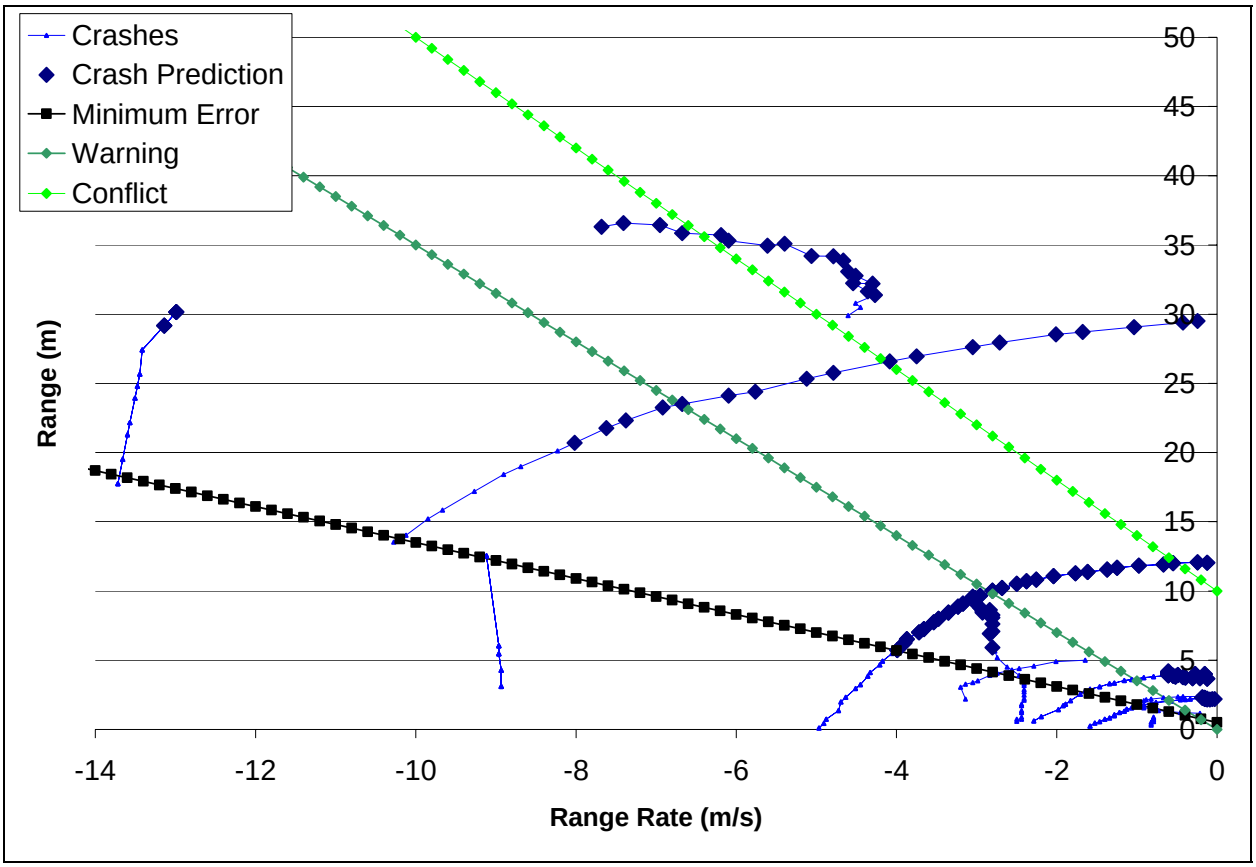


Figure 6.4. Range/range rate trajectories for all crashes. Trajectories include the data points up to 2 seconds prior to the crash. The boundaries shown are approximations of the warning and conflict boundaries used as part of the forward collision warning algorithm (Kiefer et al., 2003) and a minimum error boundary calculated for this dataset.

Figure 6.4 also provides some insight into the utility of the radar signature. If a crash occurred, one would expect that the range value of the trajectory would end at zero. However, noise in the data precluded this from happening. For example, the left most trajectory in the graph represents a case in which a driver, in order to attempt to avoid the vehicle in front of her, swerved off the road and ran into a telephone pole. She actually clipped the right corner of vehicle in front of her, but at the point of impact, the vehicle was heading to the right of the lead vehicle such that the radar on the front of the vehicle did not have the lead vehicle in its field of view. In addition, the telephone pole was not detected by the radar in a timely manner. For these reasons, the range/range rate never got to zero in the data stream.

Therefore, the limitations of currently available (and affordable) radar, in addition to complex lead-vehicle scenarios, caused considerable noise in the data. The result was somewhat unreliable data classification, regardless of the ultimate use of the data.

Attempts were also made to classify the crash trajectory data further out than 2 seconds from the ultimate crash event. As shown in Table 6.2, the ability of any of the boundary equations to accurately classify the event as a crash decreased dramatically. This finding created some difficulty with regard to the development of collision warning systems. Drivers must be

provided the warning with sufficient time to perceive the warning, assess the threat, and respond accordingly. In some lower range/higher range rate scenarios there was little time available for such a driver response.

Table 6.2. Number of lead-vehicle crashes detected by each crash boundary based upon the number of seconds prior to the crash.

	Seconds Prior To Crash				
	0	1	2	3	4
Minimum Error	10	6	0	0	0
Warning	10	9	4	2	0
Conflict	11	10	7	5	3
All in scenario	11	10	7	6	3

Following Vehicle Conflicts

For following-vehicle conflicts, there were 9 crash and 52 near-crash events comprising the 61 valid events. For analysis purposes, 157 invalid events that met the criteria discussed previously were included. Note that there were fewer following-vehicle events compared to lead-vehicle events. This was due to the differences in the radar signatures for a forward versus a rear-facing radar. Essentially, a forward-facing radar had many more objects to discern since gaining range on any static object could potentially be a threat. Alternatively, a rear-facing radar only needed to produce a signature for objects moving toward the vehicle since all other targets were increasing in range as the vehicle moved forward.

Therefore, a boundary for following-vehicle events was calculated in an attempt to minimize the overall error rate (i.e., misses + false alarms). The crash boundary for Following Vehicle conflicts was:

Minimum Error Boundary: $y = -0.9x + 0.7$

Table 6.3 provides the hit and false alarm rates for the minimum error boundary. Once again, the minimum error boundary provided some discrimination between valid and invalid events, but at the cost of missing two crashes.

Table 6.3. Percentage of hits and false alarms for each boundary model for a conflict with a following vehicle.

Conflict with Following Vehicle	Valid crash + near-crash hit rate	Invalid false alarm rate	Diff	Crash hit rate
Minimum Error	67%	33%	34%	7 / 9

Figure 6.5 represents the point of greatest threat with the Following Vehicle for the valid events and a random sample of invalid events. The boundary shown is the minimum error boundary calculated for this dataset.

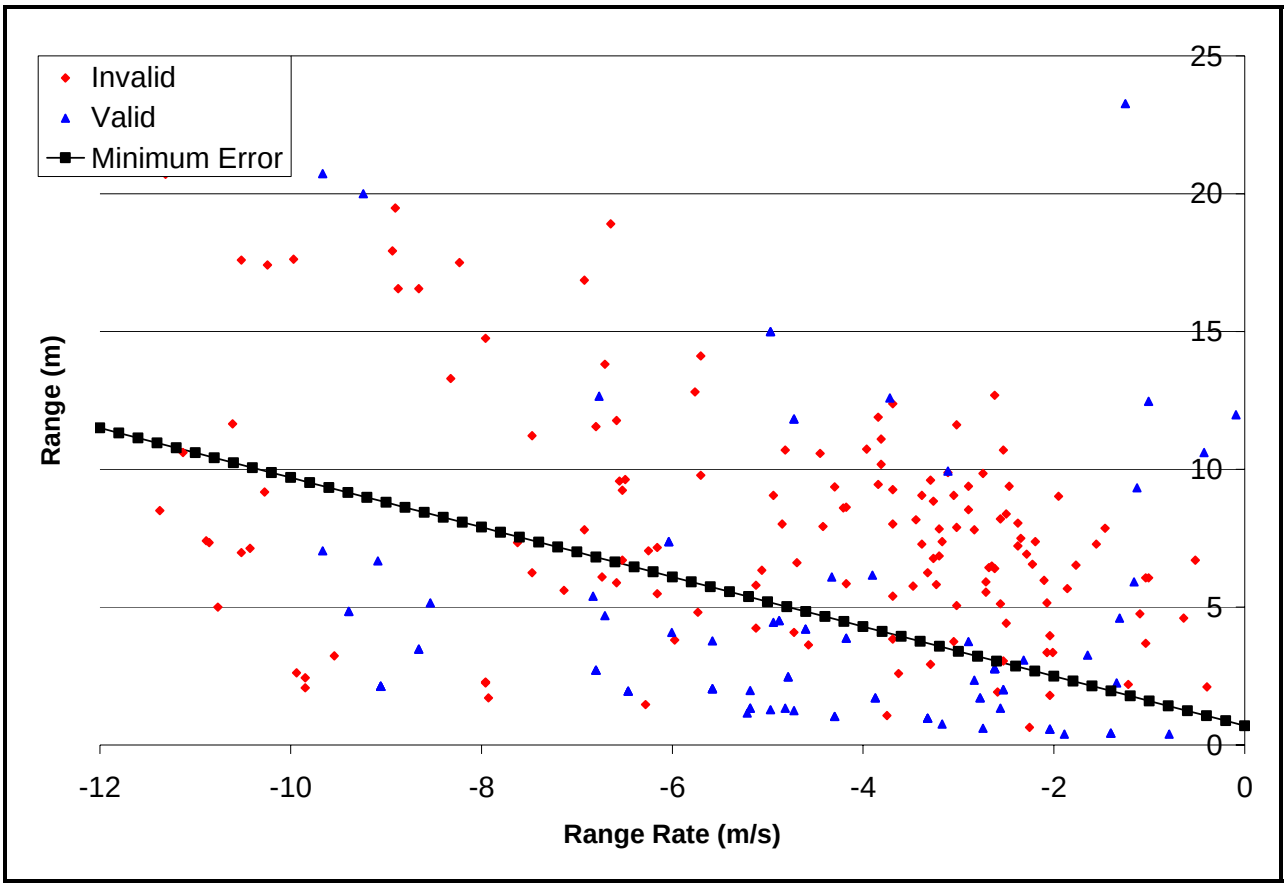


Figure 6.5. Point of greatest threat with Following Vehicle for all crashes and near-crashes, and a random sample of invalid events. The boundary shown is the minimum error boundary calculated for this dataset.

Figure 6.6 shows the trajectories of the Following Vehicle conflict data. As shown previously with Lead-Vehicle conflicts, trajectories show the timeline of a vehicle for up to 8 seconds prior to the trigger. The point at which the trajectory crossed the boundary was the point of discrimination. As before, the color blue indicates a valid event and the color red indicates an invalid event. A blue line crossing the boundary would be considered a correctly identified valid event. A red trajectory crossing a boundary would be a false alarm. Blue trajectories that do not cross a boundary are misses.

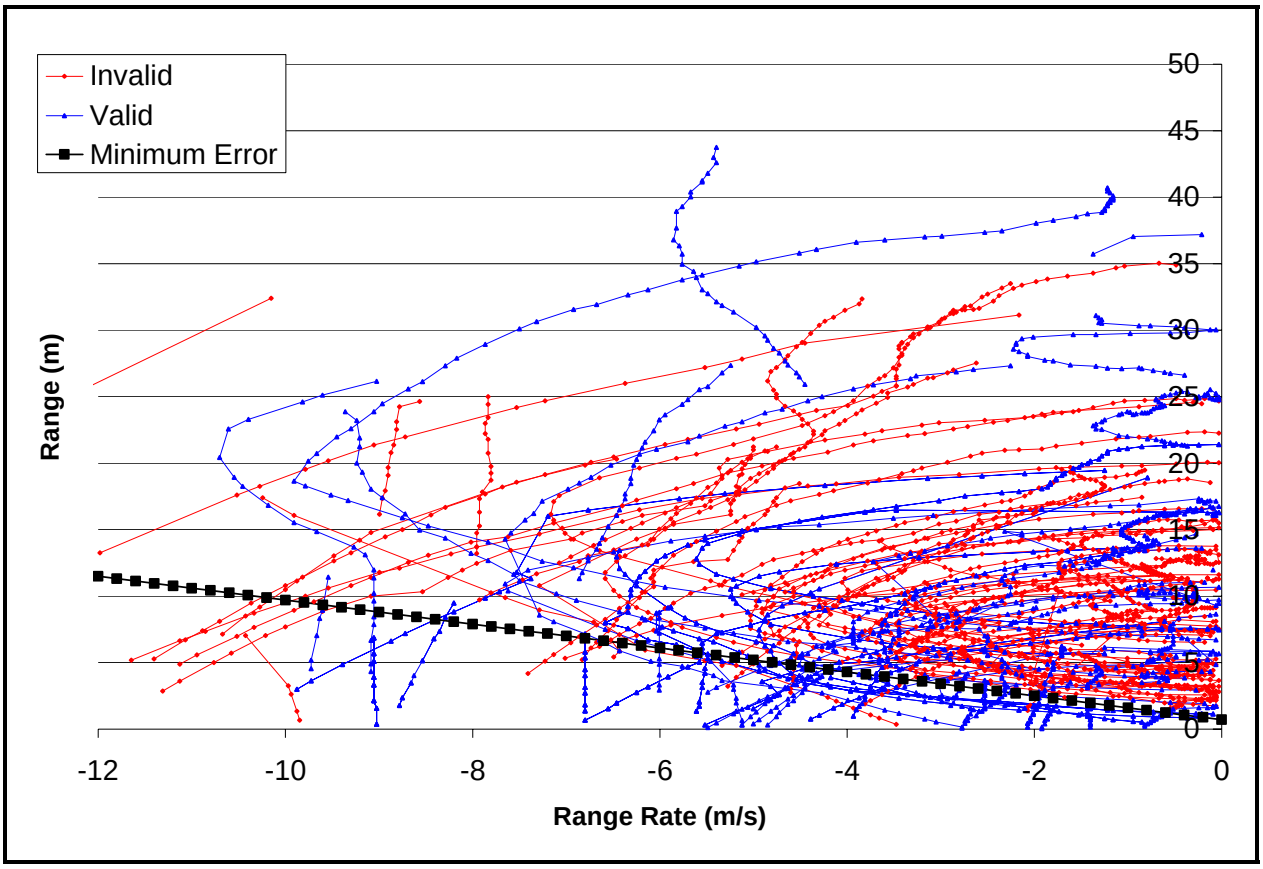


Figure 6.6. Range/range rate trajectories for following vehicles including crashes and near-crashes, and a random sample of invalid events. The boundary shown is a minimum error boundary calculated for this dataset.

To enhance the visualization, Figure 6.7 provides a random sample of the trajectories in Figure 6.6. An interesting note with regard to the signature of many of the trajectories in Figure 6.7 is that the signature of the valid and the invalid events are very similar.

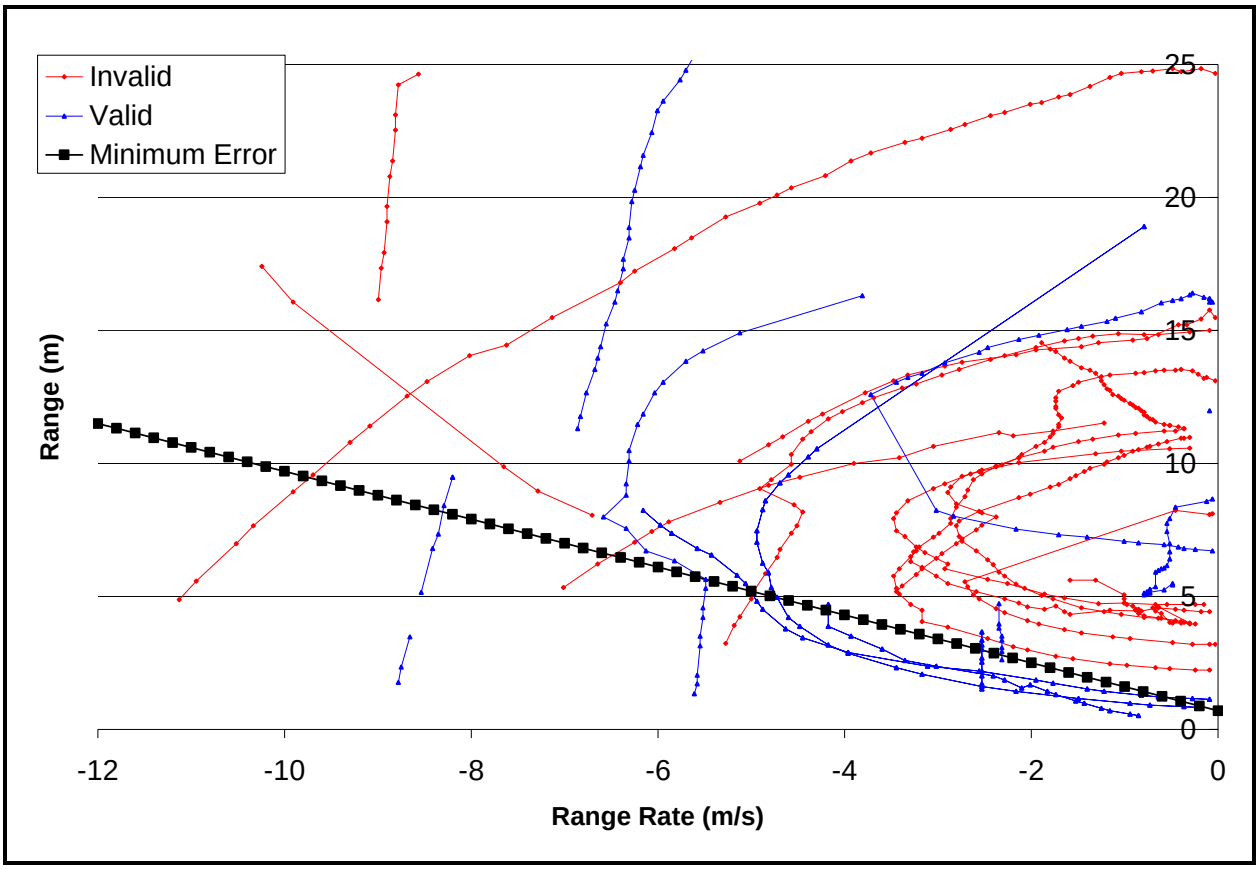


Figure 6.7. Range/range rate trajectories for following vehicles including a random sample of crashes and near-crashes, and invalid events. The boundary shown is the minimum error boundary calculated for this dataset.

As with lead-vehicle conflicts, the efficacy of providing a following-vehicle pre-crash prediction based upon the minimum error boundary is discussed. Figure 6.8 provides range/range rate trajectories for the crash events. The blue points on each trajectory represent that time during the trajectory that was at least 2 seconds prior to the collision and during which the crash could be predicted. Intuitively, one could deduce that 2 seconds is not sufficient time to evade a following vehicle that is rapidly closing in.

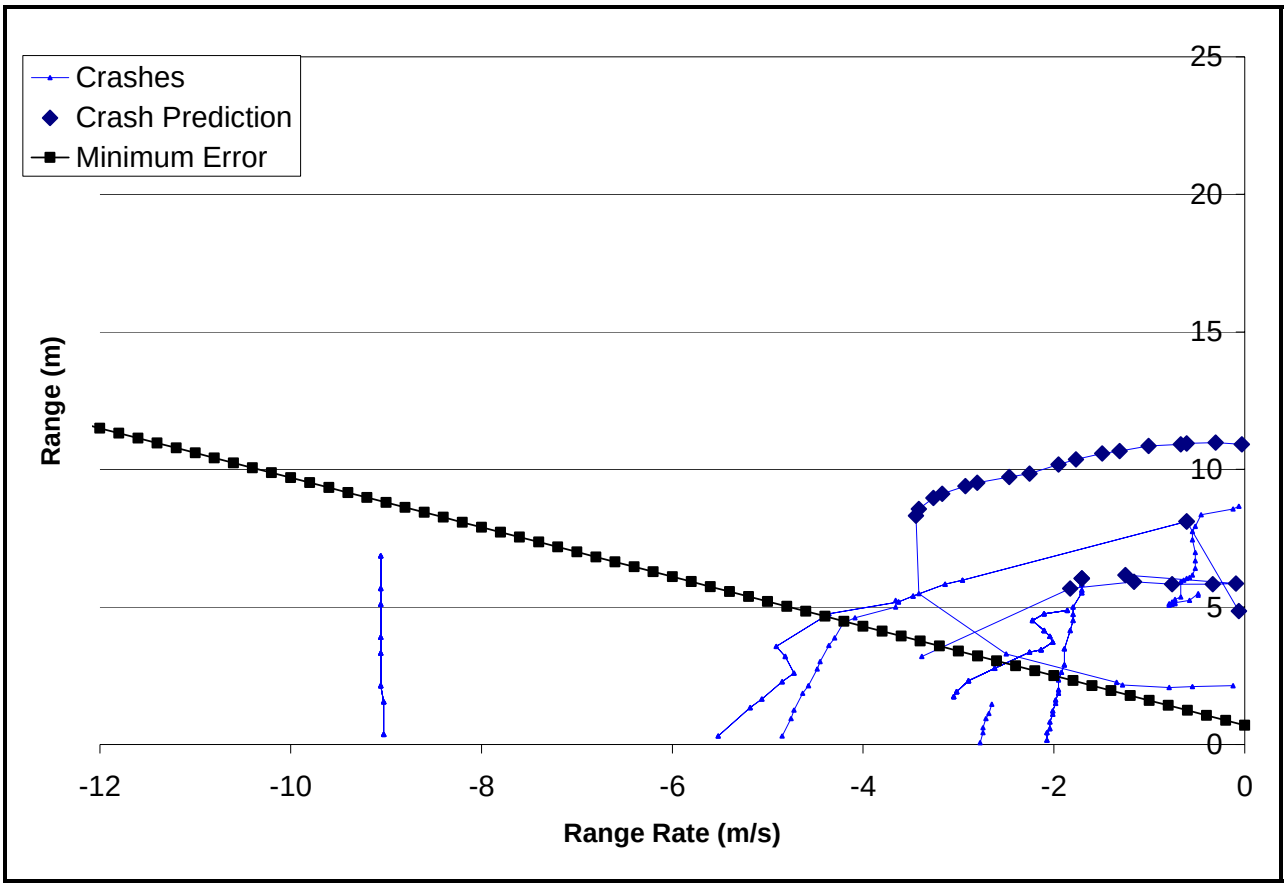


Figure 6.8. Range/range rate trajectories for all following-vehicle crash events. The crash prediction trajectory shown includes the data points up to 2 seconds prior to the crash. The boundary shown is a minimum error boundary calculated for this dataset.

Establishing a Speed Threshold

Another aspect of a quantitative crash/near-crash boundary is the role of vehicle speed on the correct classification of radar signatures. A speed threshold is often established to reduce the number of false alarms; that is, sensor data acquired below a speed criterion is disregarded. As shown in Table 6.4 the difference between correct classifications of invalid and valid events did not improve for lead-vehicle events with the inclusion of a speed threshold. In fact, these numbers generally decreased for the Minimum Error Boundary, held fairly steady for the approximated Warning threshold, and slightly improved for the Conflict boundary.

Note that the data in Table 6.4 indicate that a speed threshold greatly reduced the number of captured crash events. One issue that becomes apparent is that many rear-end crashes in the 100-Car Study occurred at low speed. While one could argue that these crashes are much less important since the likelihood of an injury is very low, they may be more important from driver acceptance, property damage, and crash-caused delay points of view.

Table 6.4. The number and percentage of events that would be correctly classified with the addition of a speed threshold for each boundary equation for lead-vehicle conflicts.

Speed Threshold by Boundary Types	Boundary Equation		Number of Events			Percent			
	-m	b	invalid	valid	crash	invalid	valid	diff	crash
Speed = 0									
Min error	1.3	0.5	1,272	224	10	20.56	74.42	53.86	90.91
Warn	3.5	0	4,540	270	10	73.39	89.70	16.31	90.91
Conflict	4	10	5,985	292	11	96.75	97.01	0.26	100.00
All in scenario	0	50	6,186	301	11	100.00	100.00	0.00	100.00
Speed = 5									
Min error	1.3	0.5	1,260	217	6	20.37	72.09	51.72	54.55
Warn	3.5	0	4,513	263	6	72.96	87.38	14.42	54.55
Conflict	4	10	5,929	287	8	95.85	95.35	-0.50	72.73
All in scenario	0	50	6,121	296	8	98.95	98.34	-0.61	72.73
Speed = 10									
Min error	1.3	0.5	1,141	200	4	18.44	66.45	48.00	36.36
Warn	3.5	0	4,280	248	4	69.19	82.39	13.20	36.36
Conflict	4	10	5,696	268	5	92.08	89.04	-3.04	45.45
All in scenario	0	50	5,890	277	5	95.22	92.03	-3.19	45.45
Speed = 15									
Min error	1.3	0.5	952	172	3	15.39	57.14	41.75	27.27
Warn	3.5	0	3,805	226	3	61.51	75.08	13.57	27.27
Conflict	4	10	5,240	250	4	84.71	83.06	-1.65	36.36
All in scenario	0	50	5,445	259	4	88.02	86.05	-1.97	36.36
Speed = 20									
Min error	1.3	0.5	752	135	3	12.16	44.85	32.69	27.27
Warn	3.5	0	3,159	194	3	51.07	64.45	13.38	27.27
Conflict	4	10	4,551	215	3	73.57	71.43	-2.14	27.27
All in scenario	0	50	4,767	226	3	77.06	75.08	-1.98	27.27

Establishing a Deceleration Threshold

Another method for potentially filtering this data is to establish a filter that removes the data for events in which the driver did not exceed a pre-specified longitudinal deceleration. The logic here is that lead-vehicle conflicts generally are associated with higher longitudinal decelerations. Therefore the filtering of lower decelerations may reduce the noise present without eliminating many of the valid events of interest. The results are shown in Table 6.5 for deceleration thresholds greater than 0.0, 0.2, 0.3, 0.4, and 0.5g. The threshold eliminated only those events in which the driver held a constant speed or accelerated throughout the event. Therefore only a limited number of cases were available for elimination..

As shown in Table 6.5, increasing the deceleration filtering threshold did reduce the number of false alarms in some cases, particularly for the minimum error threshold. In all cases, the deceleration filtering above 0.5 g reduced the number of false alarms. However, this gain was made at the cost of eliminating a number of crash and near-crash cases of interest.

Table 6.5. The number and percentage of events that would be correctly classified with the addition of a deceleration threshold for each boundary equation for lead-vehicle conflicts.

Deceleration Threshold by Boundary Types	Boundary Equation		Number of Events			Percent			
	-m	b	invalid	valid	crash	invalid	valid	diff	crash
Decel > 0g									
Min error	1.3	0.5	1,054	214	8	17.04	71.10	54.06	72.73
Warn	3.5	0	4,240	260	9	68.54	86.38	17.83	81.82
Conflict	4	10	5,640	279	10	91.17	92.69	1.52	90.91
All in scenario	0	50	5,842	288	10	94.44	95.68	1.24	90.91
Decel > 0.2g									
Min error	1.3	0.5	691	192	8	11.17	63.79	52.62	72.73
Warn	3.5	0	3,587	232	8	57.99	77.08	19.09	72.73
Conflict	4	10	4,909	248	9	79.36	82.39	3.04	81.82
All in scenario	0	50	5,093	255	9	82.33	84.72	2.39	81.82
Decel > 0.3g									
Min error	1.3	0.5	542	185	7	8.76	61.46	52.70	63.64
Warn	3.5	0	3,152	225	7	50.95	74.75	23.80	63.64
Conflict	4	10	4,388	240	8	70.93	79.73	8.80	72.73
All in scenario	0	50	4,566	245	8	73.81	81.40	7.58	72.73
Decel > 0.4g									
Min error	1.3	0.5	344	178	6	5.56	59.14	53.58	54.55
Warn	3.5	0	2,464	217	6	39.83	72.09	32.26	54.55
Conflict	4	10	3,547	229	7	57.34	76.08	18.74	63.64
All in scenario	0	50	3,713	233	7	60.02	77.41	17.39	63.64
Decel > 0.5g									
Min error	1.3	0.5	71	156	5	1.15	51.83	50.68	45.45
Warn	3.5	0	431	189	5	6.97	62.79	55.82	45.45
Conflict	4	10	588	197	5	9.51	65.45	55.94	45.45
All in scenario	0	50	620	199	5	10.02	66.11	56.09	45.45

DISCUSSION

Throughout this chapter, several reasons were noted for why the crash boundary methods did not perform perfectly. One reason was simply noise in the sensor data. In some cases, radar units missed the critical target. In the example given previously of the driver who swerved, clipped a lead vehicle, then hit a telephone pole, the lead vehicle at the point of impact and the telephone pole never appeared as targets. Alternatively, radar units detected non-critical targets, such as guardrails, when the road geometry was off angle. For these cases, more sophisticated technology and algorithms would reduce the current level of false alarms and misses.

Despite potentially correctable imperfections in the radar sensor, this data clearly showed that development of purely quantitative near-crash criteria (i.e., not requiring at least some degree of

verification by a human analyst) is not currently feasible for most cases. A primary reason for this was that the kinematic signatures associated with near-crash events were virtually identical to many common driving situations that were *not* indicative of crash risk. Thus, qualitative and quantitative criteria are dependent upon one another to some degree. A qualitative criterion must be based on quantitative criteria that are based on crash risk. Similarly, a quantitative criterion alone will not suffice without qualitative information regarding the validity of the near-crash based upon context information such as the presence of a planned versus an unplanned maneuver.

The implication for large naturalistic data collection is that to ensure proper identification of valid and invalid events there will likely be a need, at least in the foreseeable future, for video data verification of dynamically triggered events. However, as discussed in the report *Goal 10: Evaluation of the Performance of the 100-Car Naturalistic Driving Study Data Reduction Plan, Triggering Methods, and Data Analysis* (separate report), given current video technology, such verification is neither difficult nor expensive relative to the overall collection effort of such large-scale field tests. It is important to understand in reviewing these results that from a large-scale naturalistic study perspective, crash detection is reasonably straightforward since there is often a greater than 1.0g peak deceleration when the crash occurs. More problematic is the elimination of near-crash cases. However, depending on the size of the study, it may be reasonable to make an *a priori* decision to capture about of 70 percent of 25,000 or 30,000 near-crash events if the false alarm rate can be reduced to the 10 percent range. Alternatively, as will be discussed in greater detail in the *Goal 10 Report (separate report)*, the cost of a false alarm is fairly low given the capability of the data reduction tools used in this study. Specifically, a trained reductionist can sort between the presence or absence of a valid conflict using video data at the rate of about 50 per hour.

The implications of these results also highlight the difficulties for deploying forward crash warning systems in the near term. Admittedly, the analysis presented is cursory and the boundary equations simplistic. Nevertheless, the sheer number of misclassified events and the relative range/range rate position of valid and invalid events indicates that a feasible, beneficial and acceptable countermeasure system might require more sophisticated information (e.g., whether or not the driver is looking forward) or possibly braking authority instead of a simple warning.

CHAPTER 7: GOAL 3, CHARACTERIZATION OF DRIVER INATTENTION AS IT RELATES TO INCIDENTS, NEAR-CRASHES, AND CRASHES

DATA ANALYSIS OVERVIEW

Secondary task distraction and other sources of inattention have been issues in driving for many years. More recently, the increased use of cellular telephones (cell phones) and personal digital assistants (PDAs) by drivers has again raised the issue of tasks that can be safely performed in an automobile while driving. While data collected in controlled settings such as test tracks and simulators suggest that driving while performing many tasks, including cell phones, can degrade driving performance, other research suggests that driving performance when dialing a cell phone is less affected than driving performance when talking to passengers, eating, or looking for an object in the vehicle (Stutts, et al., 2003).

This chapter addresses four types of driving *inattention*, which have been operationally defined as:

- *Secondary task distraction* – driver behavior that diverts the driver's attention away from the driving task. This may include talking/listening to cell phone, eating, talking to a passenger, etc. A complete list of all secondary task distractions is provided in Appendix D.
- *Driving-related inattention to the forward roadway* – driver behavior that is directly related to the driving task but diverts driver's attention away from the forward field of view. This includes such items as checking the speedometer, checking blind spots, observing adjacent traffic prior to or during a lane change, looking for a parking spot, and checking mirrors.
- *Drowsiness* – driver behavior that included eye closures, minimal body/eye movement, repeated yawning, and/or other behaviors based upon those defined by Wierwille and Ellsworth (1994).
- *Nonspecific eyeglance away from the forward roadway* -- cases in which the driver glances, usually momentarily, away from the roadway, but at no discernable object or person. Eyeglance reduction and analysis of these cases was accomplished for crash and near-crash events only.

A two-step data reduction process was conducted to create a database for the data described above. First, the data reductionists assessed whether inattention was a contributing factor to the presence or severity of the event in question. This assessment required the presence of two separate criteria: (1) The reductionists looked for instances in which the presence of driver inattention occurred *within 3 seconds* of the onset of the conflict or at the onset of the conflict, and (2) The reductionists assessed whether the presence of the inattention contributed to the presence or severity of the event. This was accomplished by assessing factors such as driver reaction time to determine whether the driver's initial performance or subsequent response was consistent with the inattention in question. Examples of some of the most common situations include: an inopportune glance away from the roadway to check a blind spot at the precise moment of an unexpected forward event, an inappropriate level of secondary task engagement

leading to a lane tracking error in the presence of other traffic, or a case of drowsiness resulting in a delayed reaction time.

A second data reduction process occurred using eyeglance analysis for all near-crash and crash events. For this analysis, eyeglances away from the forward roadway were reduced to determine driver eyeglance location and duration. These cases also included inattention analysis regardless of whether inattention was judged to be a contributing factor in the analysis described above. This was accomplished because there were a number of cases of short glances away from the forward roadway that probably contributed to the presence or severity of events that were not apparent during the initial review process. For this analysis, the last eyeglance location away from the forward roadway that occurred during a window of 3 seconds prior and 1 second after the onset of the conflict was used for classification. If the eyeglance was to a specified location (rear-view mirror, toward passenger, or toward cell phone), the eyeglance is categorized as either driving-related inattention or secondary task distraction as defined above. If the eyeglance is toward an internal or external location and the source is unknown, then the eyeglance is categorized as "non specific."

Driver Data Included in the Analyses

As discussed in Chapter 2, *Method*, 109 primary drivers were recruited to participate in this study; however, data reduction was effectively conducted on 241 participants since many of the primary drivers allowed family members and friends to drive their vehicles. There was no unobtrusive, feasible method of determining driver identification in the raw data, so driver identification was performed during data reduction. Demographic information for the additional drivers was not obtained.

For data analyses that do not consider driver age, gender, or vehicle miles traveled (VMT), data from all 241 participants was used in the analyses. For data analyses that included age and gender, the 109 primary drivers were used (those with fewer than 1,000 VMT were excluded, leaving 98 drivers in the analyses).

VMT per driver was estimated only for primary drivers. Estimates were calculated based on video reduction during which reductionists viewed a sample of 100 trip files for each vehicle and recorded whether the primary driver was operating the vehicle. The proportion of trip files that the primary driver was behind the wheel was multiplied by the total VMT for that vehicle to arrive at a VMT estimate for each primary driver. VMT was not calculated for secondary drivers since speed sensor data for this group indicated that they drove fewer than 1,000 miles for the year.

Overall Rate of Events by Driver

Figure 7.1 provides an overview of the total number of events per driver per vehicle mile traveled. Rate is calculated with the frequency of events for each driver divided by their VMT. To obtain the rate of events per million vehicle miles traveled (MVMT), the total of events was multiplied by 1,000,000. As shown in Figure 7.1, the rate of events per driver was highly variable. It is important to consider this variability in considering the analyses described in this section.

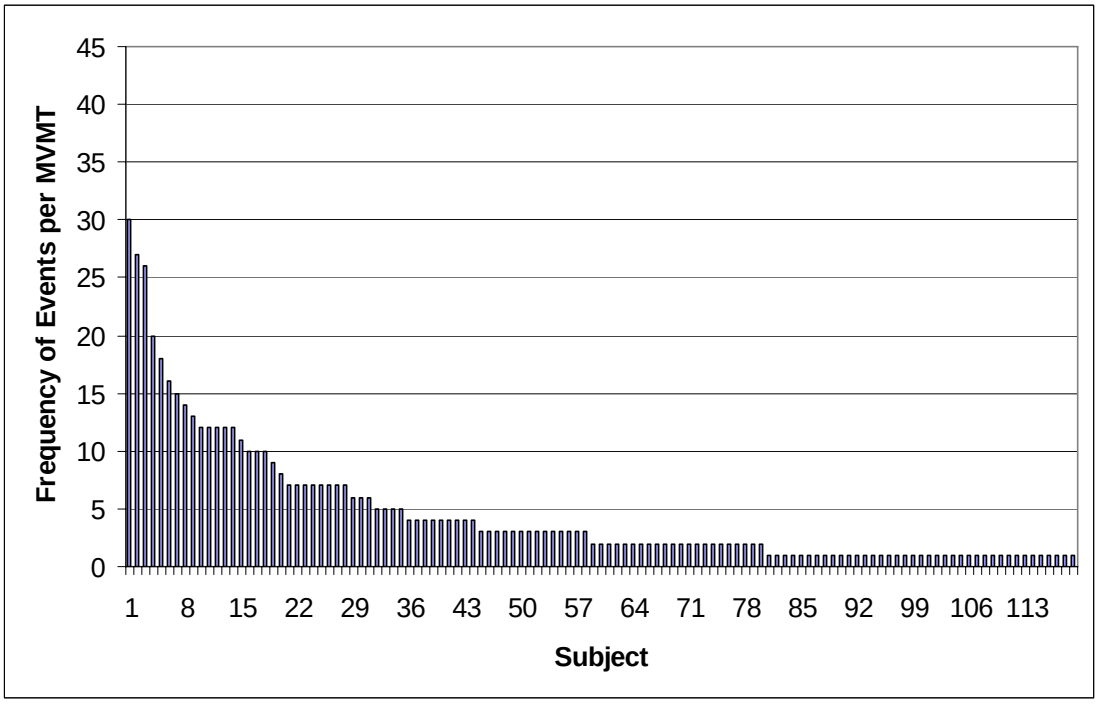


Figure 7.1. Number of events per MVMT (N=98).

Figure 7.2 below shows the subject variability in the rate of inattention-related event occurrences. In each of the events below, the driver was either labeled by data reductionists to be engaging in a secondary task, inattentive to the forward roadway, drowsy, or looked away from the forward roadway at a nonspecific object or person (for crashes and near-crashes only). The range in variability, while still very high, is somewhat reduced from the overall subject variability rate of occurrence of total events.

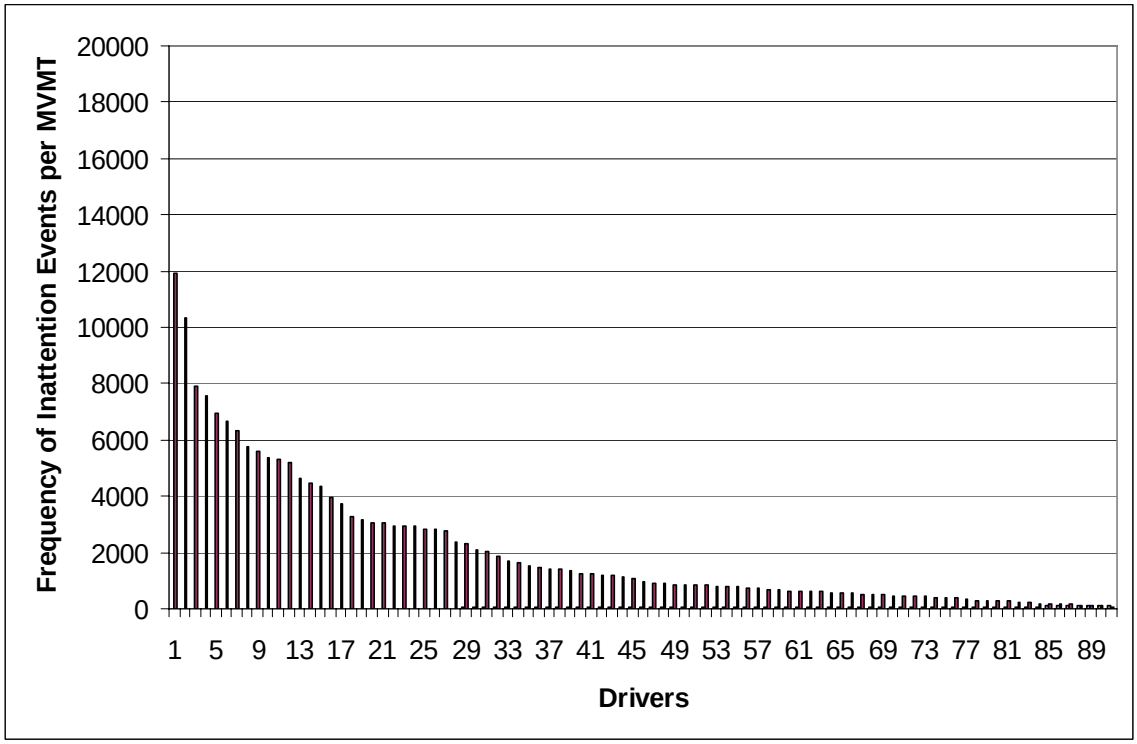


Figure 7.2. Frequency of inattention-related events per MVMT in which inattention is due to: (1) drowsiness, (2) inattention to the forward roadway, (3) secondary task, (4) specific eyeglance away from forward roadway (for crashes and near-crashes only) or (5) nonspecific eyeglance away from forward roadway (for crashes and near-crashes only).

Question 1. What is the relative frequency of events for which driver inattention was a contributing factor? What is the relative frequency of occurrence of driver inattention events versus non-driver inattention events for incidents, near-crashes, and crashes?

Figure 7.3 shows the total frequency of crashes, near-crashes, and incidents for which drivers were inattentive versus those events for which the driver was attentive. Note that the frequency of incidents is two orders of magnitude higher than the frequency of crashes; percentages of total crashes and near-crashes will be used when appropriate for the remainder of this chapter to aid in readability.

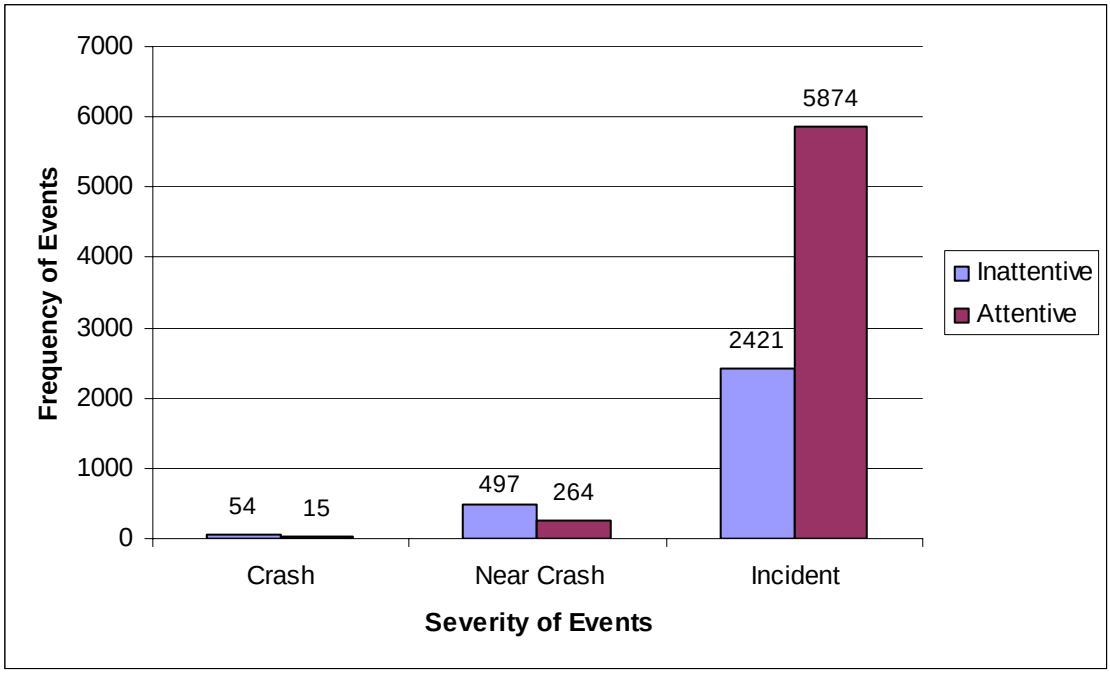


Figure 7.3. Frequency of crashes, near-crashes, and incidents for which drivers were inattentive versus attentive (Driver N = 241).

Percentage of Attentive versus Inattentive-Related Events by Severity Level
It was determined that the relative frequency of events for which driver inattention was a contributing factor versus the relative frequency of events for which driver inattention was not a contributing factor. Inattentive events included those cases for which the reductionists identified the driver as being in one of the four categories of inattention: secondary task; driver-related inattention to the forward roadway; drowsiness; or nonspecific eyeglance away from the forward roadway (crashes and near-crashes only). Attentive drivers were not engaged in these behaviors.

The percentage of events of differing severities identified in the 100-Car Study database as having driver inattention listed as a contributing factor is shown in Figure 7.4 (i.e., The crashes that were marked as inattentive *plus* the crashes that were marked attentive is *equal to* total number of crashes). The overall percentage of driver inattention-related events decreased with

decreasing event severity. For the least severe, *incident* category, the majority of events (i.e., 71%) did not involve inattention. An important finding that will be discussed in several sections of this report is that the most severe events generally had multiple factors associated with them. As indicated in Figure 7.4, the majority of the crash and near-crash cases involved a combination of factors that included a precipitating event (which was commonly present) in conjunction with some form of driver inattention to the forward roadway. Conversely, incidents often occurred after a precipitating event, but while the driver was attentive. Thus, these results indicate that inattention often leads to *increased* incident severity, leading to more near-crash and crash circumstances. It also indicates that there is at least one important conceptual difference between incidents versus near-crashes/crashes. Specifically, incidents may not be as predictive of the combinations of factors that lead to crashes, but may be more predictive of the presence of precipitating events.

An important finding of this report is that almost 80 percent of all crashes and 65 percent of all near-crashes involved the driver looking away from the forward roadway just prior to the onset of the conflict. Further analyses of eyeglances for incidents are underway and will be reported in a follow-up report.

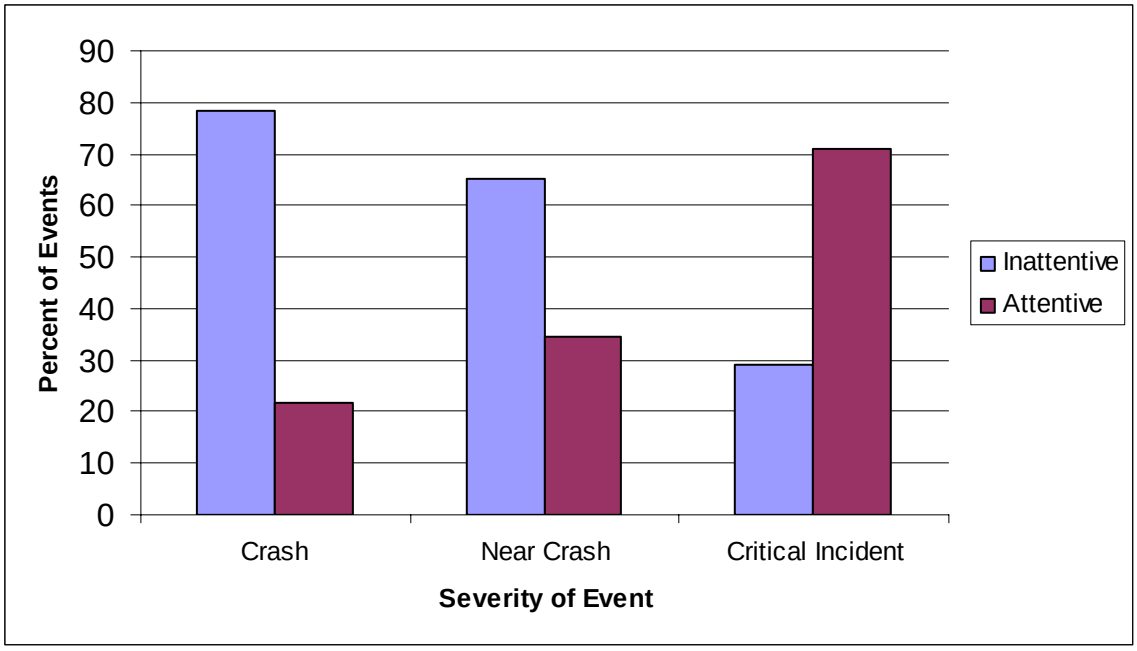


Figure 7.4. Percentage of events that drivers were inattentive versus attentive by severity level (Driver N = 241).

Percentage of Events for which Inattention was a Contributing Factor

As stated previously, the majority of crashes and near-crashes identified in the 100-Car Study database had at least one type of driver inattention listed as a contributing factor. Obtaining the frequency of events was conducted by calculating the frequency of each category of inattention as well as each combination of attention types as there were many crashes, near-crashes, and

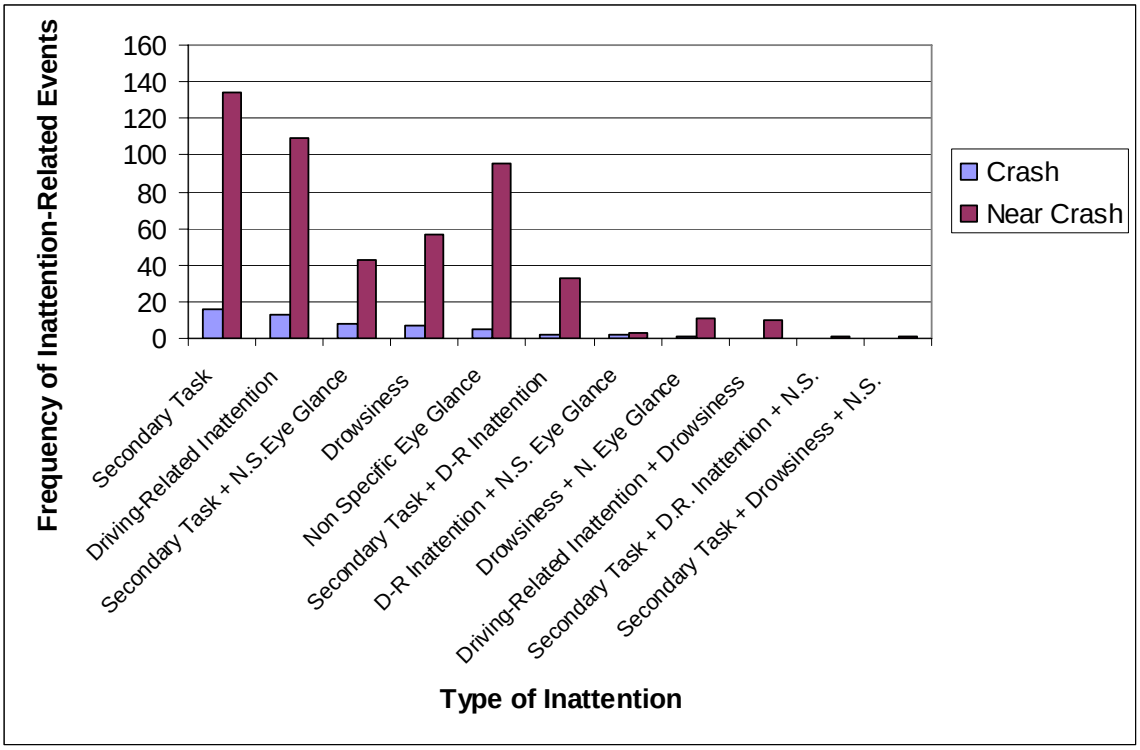
incidents that contained more than one type of inattention (e.g., one crash had drowsiness listed as a contributing factor as well as driving-related inattention to the forward roadway). Given that this is the first dataset where it is possible to determine a single or multiple sources of driving inattention, all combinations will be presented and discussed.

Figure 7.5 shows the frequency of each inattention type for crashes and near-crashes. Note that combinations of factors are also presented. As shown, secondary task distraction was associated with the highest percentage of crash and near-crash events followed by driving-related inattention to the forward roadway. It is important to note that these data represent percentages derived from raw frequencies, thus exposure is not accounted for. Specifically, one needs to determine the frequency and duration with which each of these categories of inattention are present during normal, non-event, driving in order to make judgments about relative risk. As of this writing, an additional analysis is underway that will establish exposure and make such relative risk comparisons.

Figure 7.5 and Figure 7.6 also show that for the inattention-related events, that there is a significant component of drowsiness. Over 10 percent of the inattention-related crashes and near-crashes had driver drowsiness as a contributing factor. An interesting additional observation for these cases is that the majority of the events occurred during the day, many during the morning commute.

Nonspecific eyeglances also contributed to a relatively high number of crash and near-crashes. Figure 7.6 shows that almost 10 percent of the inattention-related crashes and almost 20 percent of the inattention-related near-crashes involved cases where the driver looked away from the forward roadway, but not to any apparent location.

An important aspect of the data shown in Figure 7.6 is the degree to which the crash percentages mirror the near-crash percentages. In fact the three highest single-case categories, secondary task distraction, driving-related inattention and drowsiness, are all very close in this regard. The case where crash and near-crash percentages differ the most is the nonspecific eyeglance category (Figure 7.6). One possibility is that the level of inattention associated with this category is somewhat lower than some of the other inattention categories (either singly or in combination) perhaps resulting in a higher percentage of successful evasive maneuvers.



*The values for these types of inattention were < 1.0 percent.

Figure 7.5. The frequency of crashes and near-crashes in which these types of inattention were identified as a contributing factor.

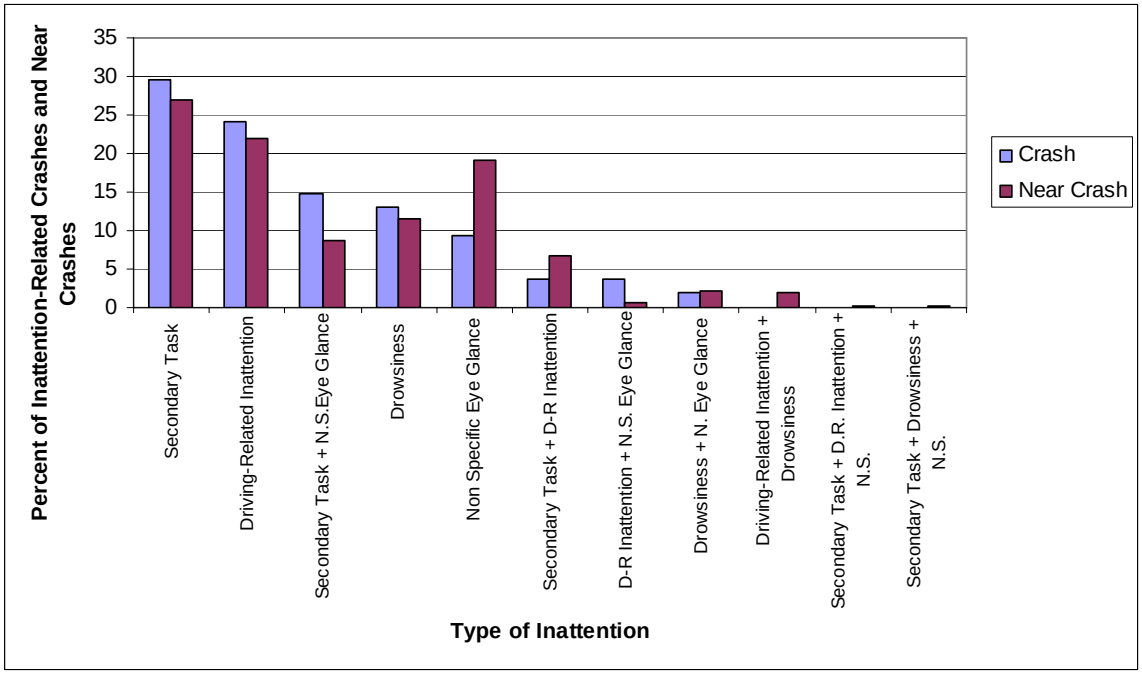


Figure 7.6. The percentage of crashes and near-crashes in which these types of inattention were identified as a contributing factor.

Figure 7.7 shows the breakdown of the types of inattention to the forward roadway for incidents. As was discussed previously, no eyeglance analysis was conducted for the incidents as part of this study, so the nonspecific eyeglance categories are not provided. Note that the majority of the inattention-related incidents involved secondary task engagement, followed by drowsiness, and driving-related inattention to the forward roadway.

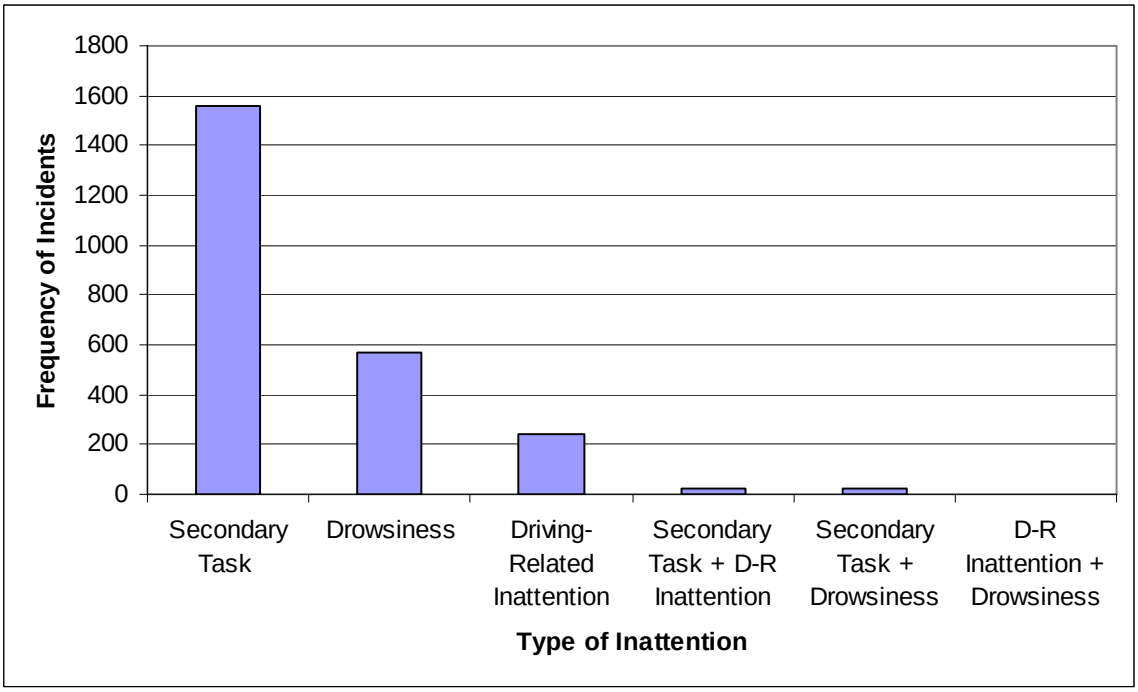


Figure 7.7. The frequency of incidents in which these types of inattention were identified as a contributing factor.

Rate of Events for Severity Level by Attention
The number of events for each driver was divided by their VMT to account for exposure. This number was then averaged across drivers and multiplied by 1,000,000 to determine rates per MVMT.

Please note that the rate of inattention-related events is higher for both crashes and near-crashes, but as discussed previously, only 29 percent of the incidents have inattention listed as a contributing factor. Again, the nonspecific eyeglance data is not included in the incident category since these data were not reduced as part of the scope of this project. However, it is hypothesized that the number of incidents with inattention as a contributing factor will be higher once it is possible to include these events. Regardless, the results shown in Figure 7.8 demonstrate that inattention affects a higher proportion of crashes and near-crashes than incidents.

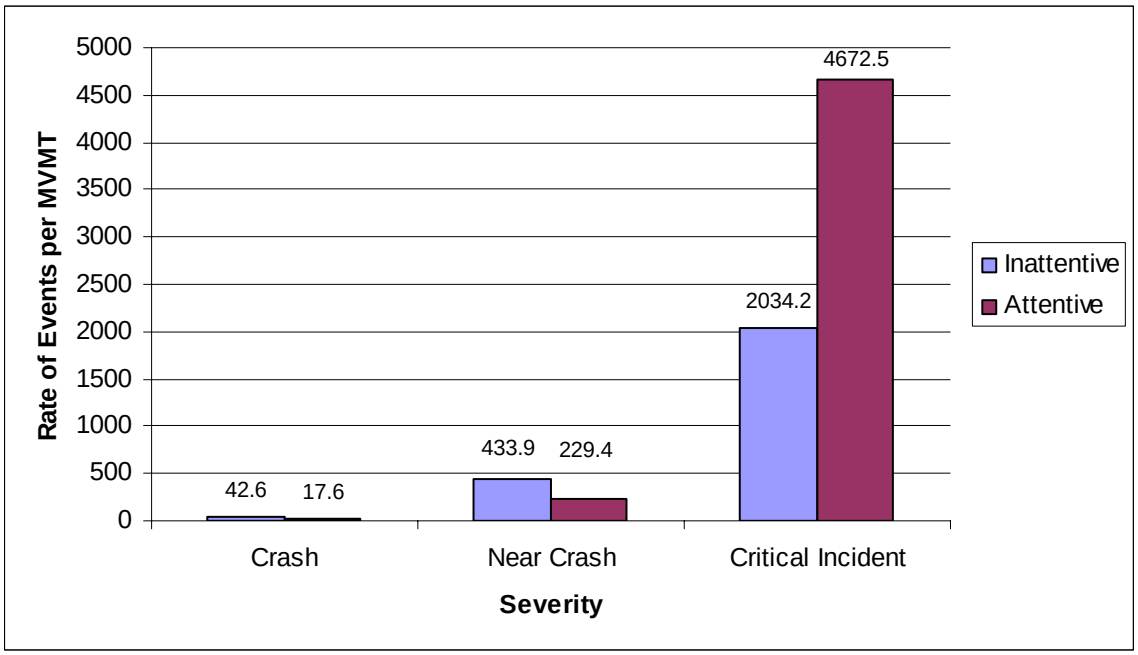


Figure 7.8. Comparison of inattention and attention-related events by severity level per MVMT (Driver N = 98).

Rate of Events per MVMT for Age by Attention

The rate of events for each age group was calculated in a similar manner as the rate of events by each level of severity level (Figure 7.9). The rate of attentive events is higher for all age groups, however, the overall rate of event occurrence as well as the rate of occurrence of inattentive events is significantly higher for the 18- to 20-year-old age group than for any other age group, $F(5,91) = 4.44$, $p < 0.01$. This finding is not surprising as it is a well-documented finding that younger drivers are involved in more crashes than are in age groups. It is also not surprising that the younger age group was involved in more inattentive events as Stutts et al. (2003) reported that younger drivers were more distracted by participants in the vehicle than other age groups as well as by other types of secondary tasks.

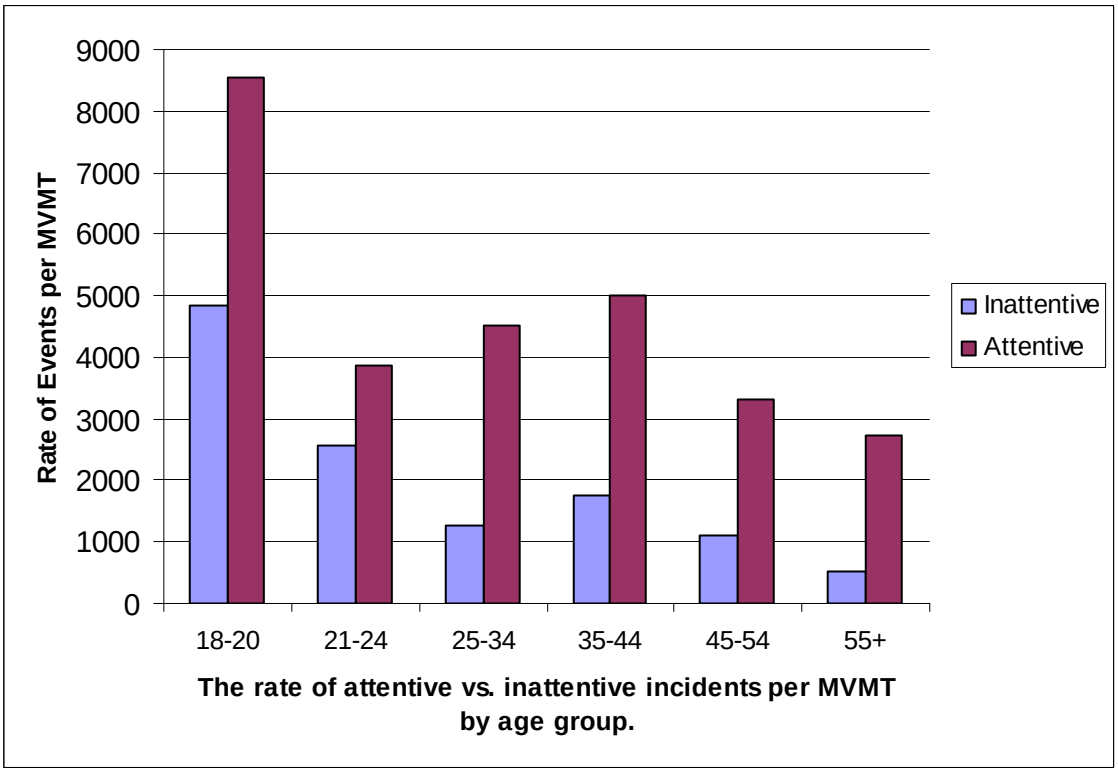


Figure 7.9. The rate of inattentive versus attentive events per MVMT by age group.

The rates of attentive versus inattentive drivers by age are also shown for crash (Figure 7.10) and near-crash (Figure 7.11) events. As shown, the rate of inattention-related crash and near-crash events decreases dramatically with age, with the rate being as much as four times higher for the 18- to 20-year-old age group relative to some of the older driver groups. This again supports the need to develop countermeasures that limit distractions and perhaps educate younger drivers of the hazards associated with inattention to the forward roadway form all of the sources shown in the section.

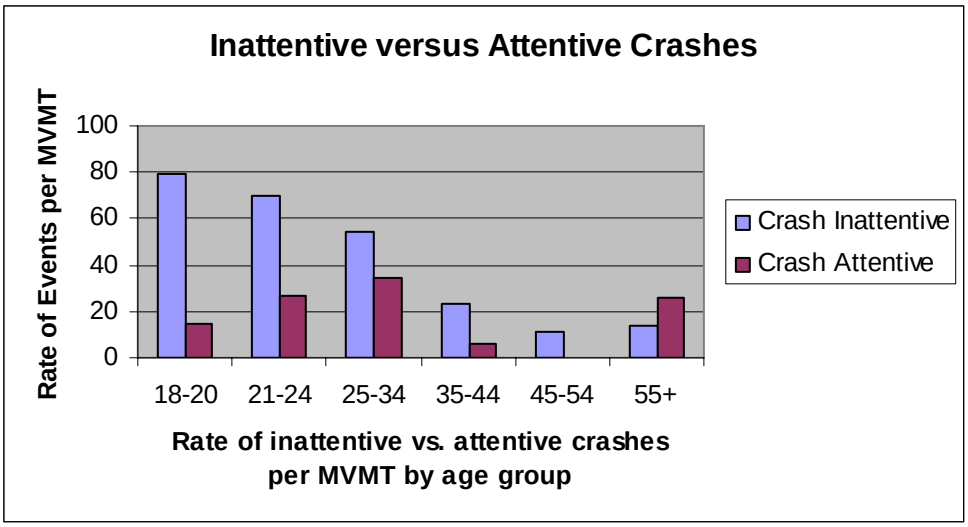


Figure 7.10. The rate of inattentive versus attentive crashes per MVMT by age group.

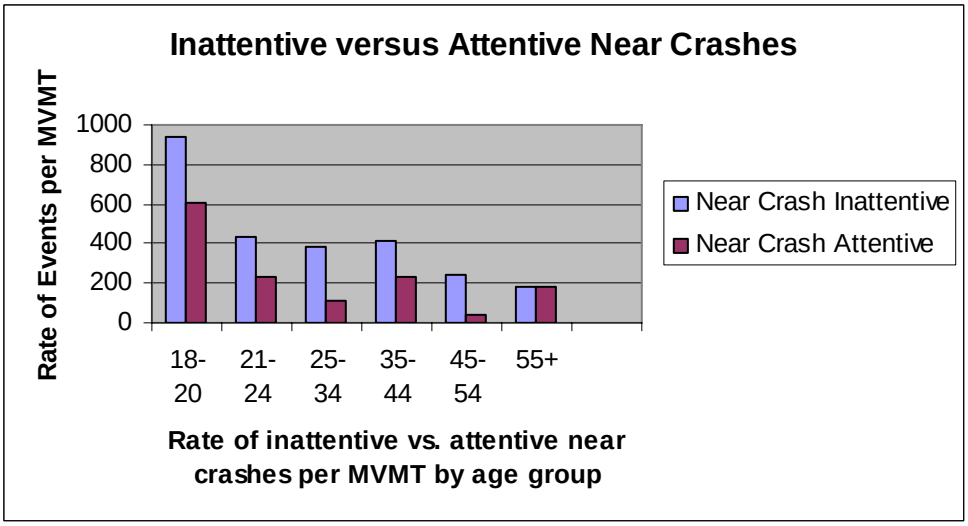


Figure 7.11. The rate of inattentive versus attentive near-crashes per MVMT by age group.

Question 2. What is the relative frequency of types of inattention involved in incidents, near-crashes, and crashes? Do some types of inattention result in more severe driving events?

To address Question 2, the frequencies and percentages will be presented for "occurrences" instead of by "events." As was shown in the prior section, more than one category of inattention was sometimes classified for a single event. To account for this, we included the inattention classification in both categories instead of presenting combinations (i.e., for readability due to the number of categories) or prioritizing one classification over another. Thus, the total number of occurrences, as depicted in the following figures, will exceed the total number of events depicted in the figures presented up to this point.

The analyses for Question 2 used the data from all of the drivers (driver N = 241). Figure 7.12 shows the frequency of each type of secondary task inattention-related occurrence. Note that *wireless devices*, including primarily cell phones and personal digital assistants (PDAs), account for the highest frequency of inattention-related occurrences, while passenger-related inattention was the next most frequent secondary task.

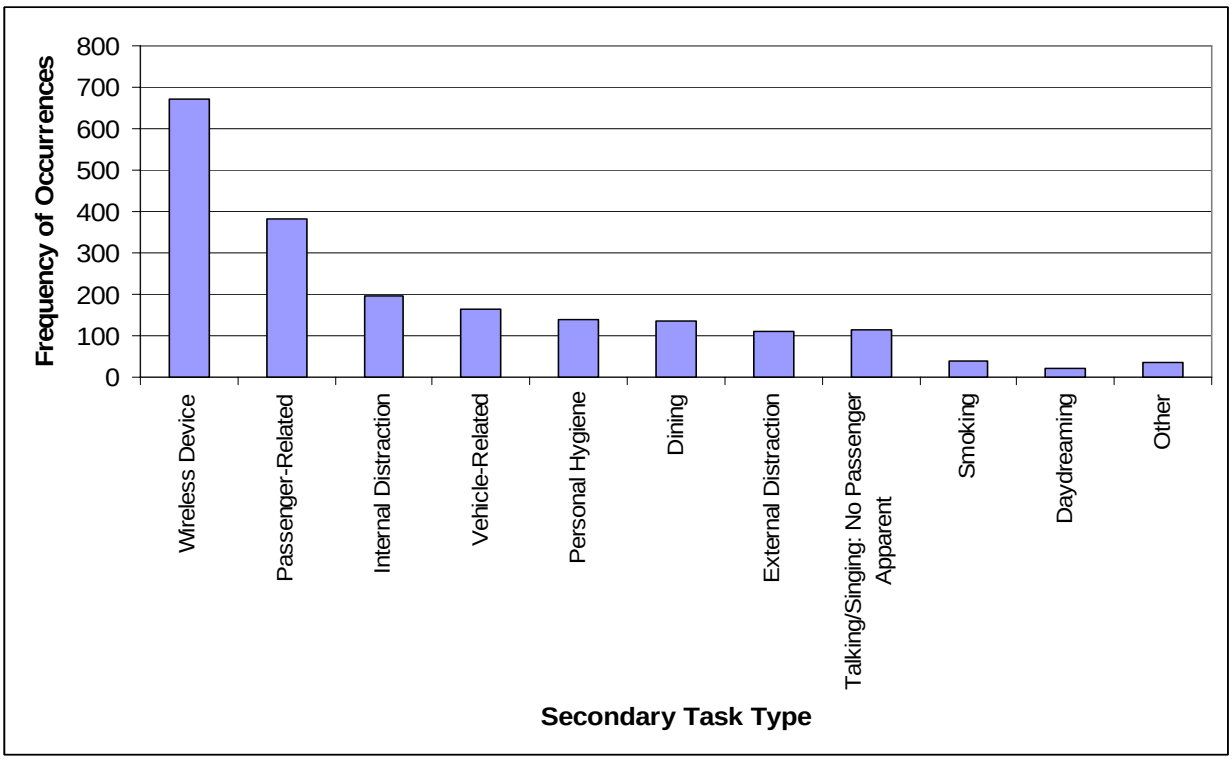


Figure 7.12. Comparison of the number of occurrences of the presence distracting agent as a contributing factor (Driver N = 241).

Figure 7.13 shows the secondary task distraction types broken out for crash and near-crash events. As shown, the most frequent secondary tasks contributing to crashes were *internal distractions*, *wireless devices*, and *passengers*. The most frequent types of inattention for near-crashes and incidents were *wireless devices* and *passenger-related tasks*.

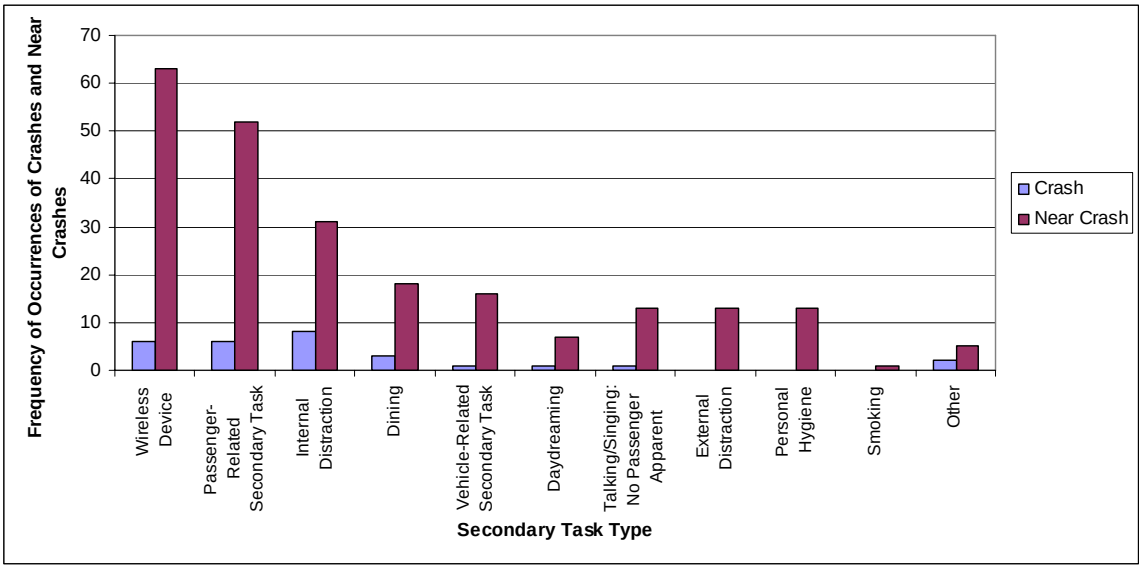


Figure 7.13. Comparison of crashes and near-crashes the frequency of occurrences of the presence, either alone or in combination, of the distracting agent as a contributing factor (Driver N = 241).

Figure 7.14 shows the frequency of driving-related inattention occurrences involving glances away from the forward roadway for each level of severity. *Left window* and *right window* glances were the most frequent contributors to the total number of inattention-related incidents. However, *center mirror* glances occurred most frequently during near-crash events and *left* window glances occurred the most during driving-related inattention crash events. As was stated previously, it is important to note that these numbers represent raw frequencies and do not account for exposure in terms of frequency of glances in non-event driving.

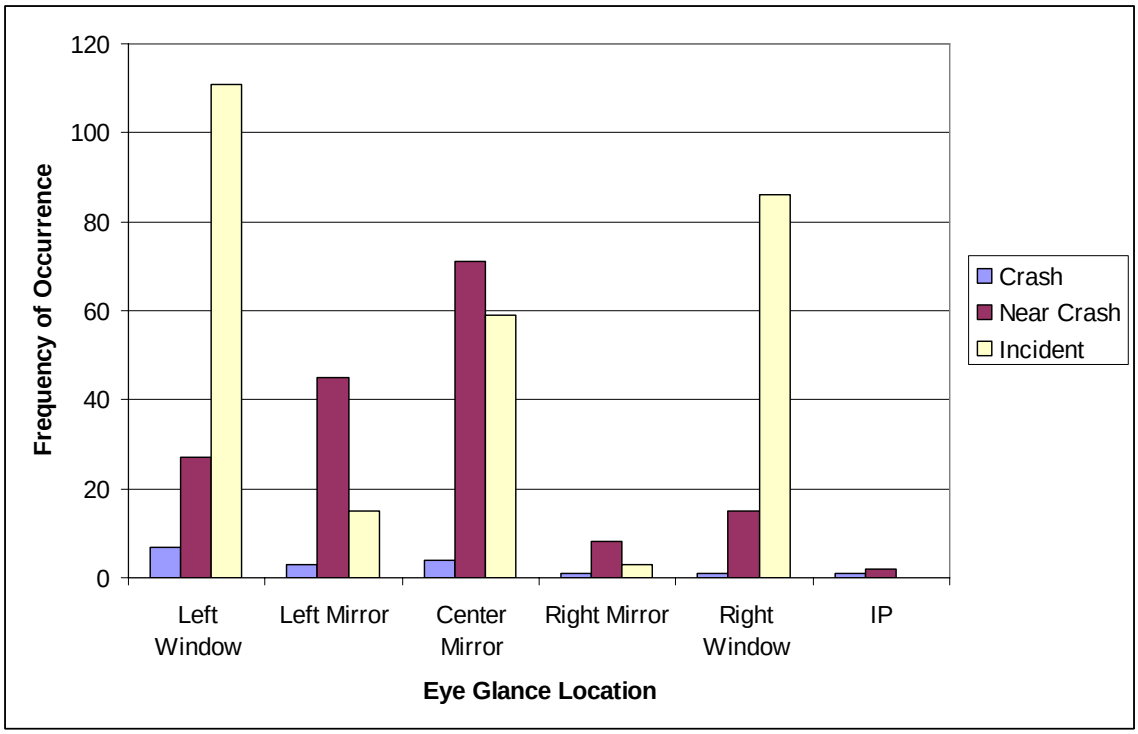


Figure 7.14. Frequency of occurrence of driving-related inattention to forward roadway, alone or in combination, by level of severity (Driver N = 241).

Frequency of Secondary Tasks

Figures depicting the numbers of occurrences for each type of secondary task are presented in the following section. Please note that a few of the figures may exceed the y-axis. These bars contain an "explosion mark" with the frequency shown in numerical form. This was done to improve the readability of some of the lower frequency occurrences.

Wireless Devices. Figure 7.15 shows a frequency distribution of all the wireless device tasks recorded by the data reductionists. The categories were all defined by name except for the *cell phone – other* category. This category was added to include all the events for which the driver was clearly not dialing or talking on the cell phone but rather looking at the display as if screening phone calls or reading text messages. *Talking/listening* on a cell phone was most frequently cited as a contributing factor to conflicts while *dialing or answering* the cell phone was the second most frequent contributing factor. While this finding could be an artifact of task duration or exposure, a considerable body of research suggests that dialing degrades driving performance more than talking or having a conversation (Jenness, et al., 2002). While this may be true, these results suggest that drivers are involved in more traffic conflicts while engaged in cell phone conversations than while dialing.

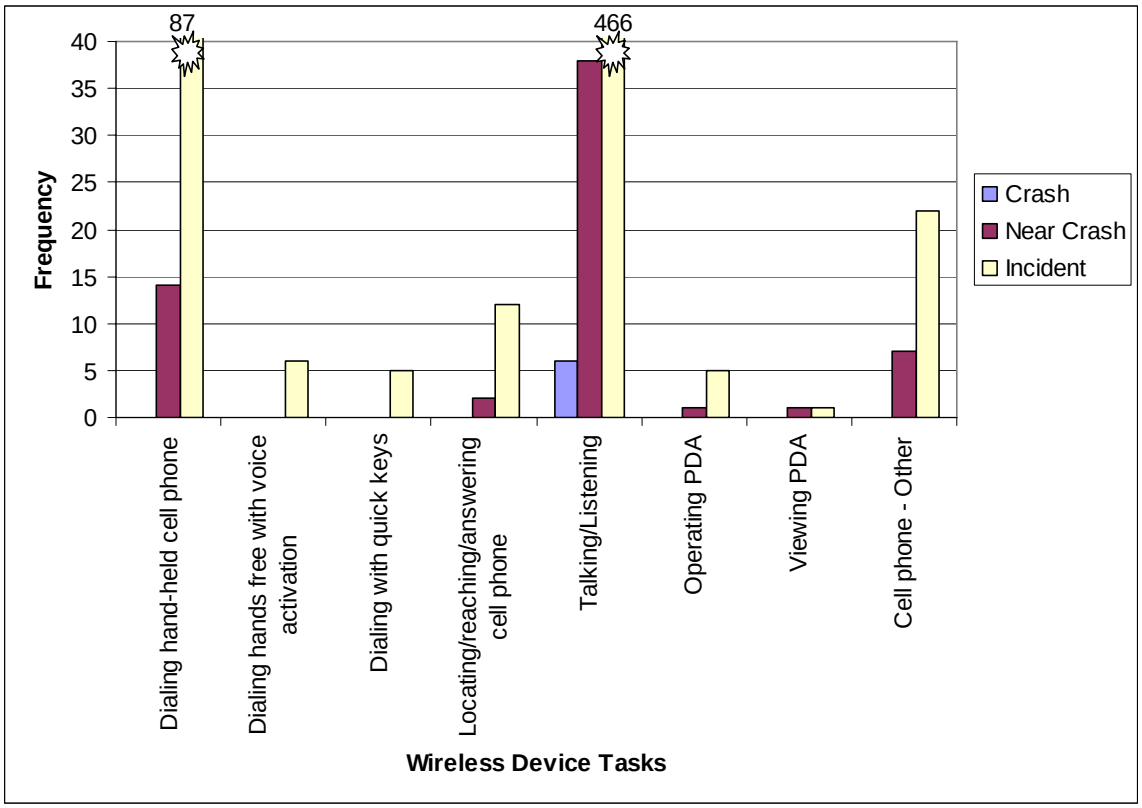


Figure 7.15. Frequency of occurrences in which the contributing factor was wireless device use (Driver N = 241).

In-Vehicle-System-Related Secondary Tasks. Figure 7.16 shows the distribution of in-vehicle-system-related secondary tasks. This category was a compilation of events directly related to the operation of in-vehicle system devices (i.e., radio or HVAC system). The category *adjusting other in-vehicle system operations* included events for which the reductionists could either: (1) not distinguish whether the driver was adjusting either the radio or the HVAC system; or (2) the driver was adjusting an added in-vehicle system component. Drivers were involved in more near-crashes and incidents while *adjusting the radio* than any other in-vehicle system operation, including inserting cassettes or compact discs. Only one crash occurred while the driver was either adjusting the radio or the HVAC system.

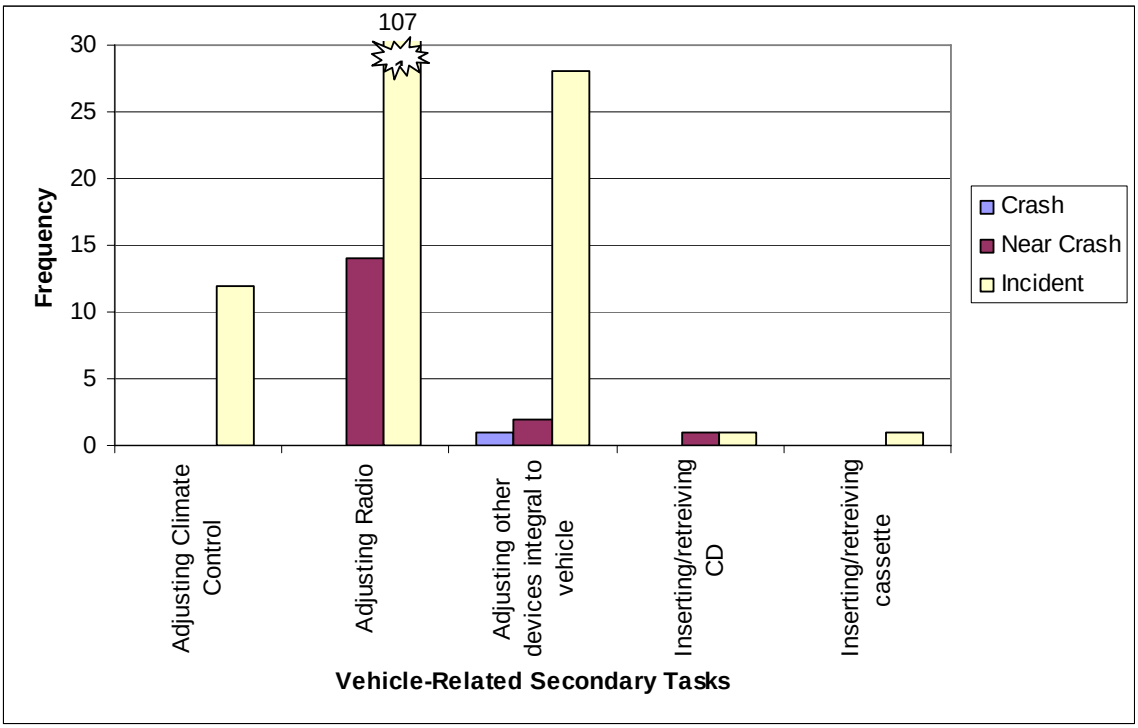


Figure 7.16. Frequency of occurrences for which the contributing factor was a vehicle-related task.

Passenger-Related Secondary Tasks. As shown previously (figures 7.12 and 7.13), passenger-related secondary tasks were the second most frequent cause of inattention associated with crashes, near-crashes, and incidents. The breakdown of the type (i.e., adult or child) and location of the passengers is shown in Figure 7.17. While it was somewhat difficult to ascertain, due to camera views, whether there were passengers, it was often possible to observe a hand reach across the camera views or view a leg in the passenger's seat in the over-the-shoulder camera view. It was also possible at times to identify the strong possibility of a passenger if the driver was clearly speaking/gesturing and looking towards the passenger seat. When this was possible, the reductionists marked *passenger in the adjacent seat* as the cause of inattention. If the driver was clearly vocalizing but did not frequently glance over at the passenger seat or in the rear seat of the vehicle, reductionists would mark this *talking/singing*. Therefore, in the *talking/singing* category, it is unknown whether or not there was a passenger in the vehicle. It was sometimes difficult to determine if there were children in the back seat, so some of the general *talking/singing* may fall into this category.

The frequency of events for which there was a passenger in the adjacent seat was by far the most common for all levels of severity. Other studies have found that children in the rear passenger seats are a frequent distracter (Stutts, et al., 2003). Since many of the subjects in this study were younger and because it was difficult to ascertain whether a child was in the rear seat, these factors may have contributed to the low occurrence of events for this category.

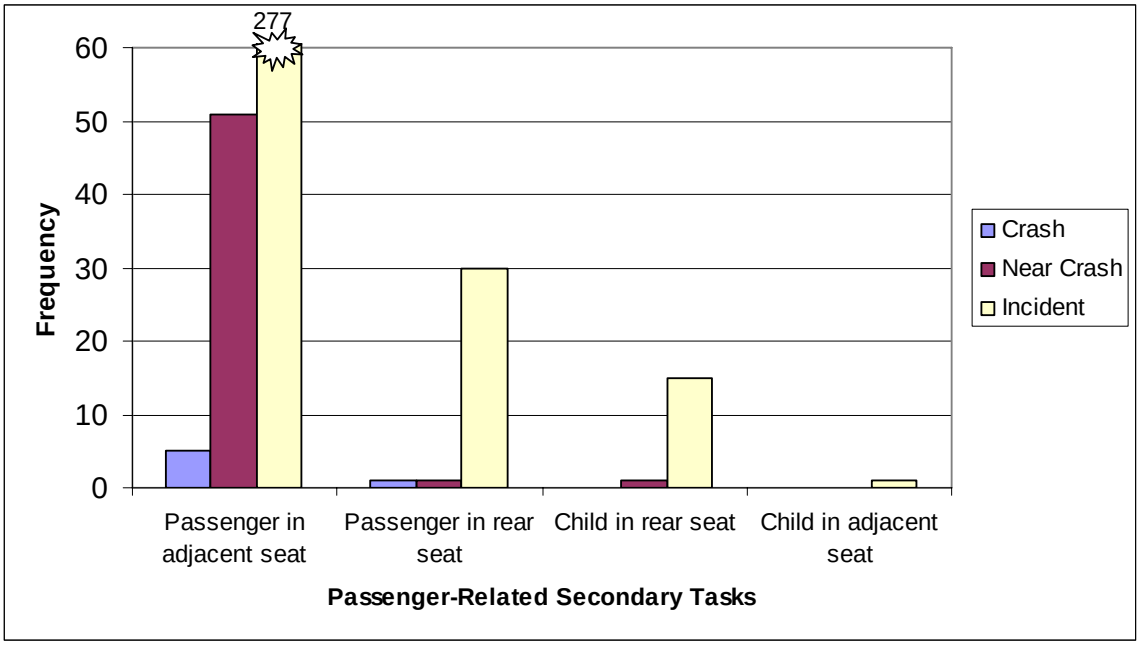


Figure 7.17. Frequency of occurrences in which the contributing factor was passenger-related secondary tasks. (Driver N = 241).

There was only one crash that occurred while the driver was vocalizing with no apparent passenger present, however there were 100 incidents and 12 near-crashes that occurred (Figure 7.18). Again, while a passenger may have been present in these events as well, it was sometimes

difficult to ascertain this during reduction. Please recall, as discussed in the Chapter *2: Method for Phase II Field Test* that due to the IRB and Certificate of Confidentiality requirements, no other passengers could be captured on video to maintain the anonymity of uninformed participants.

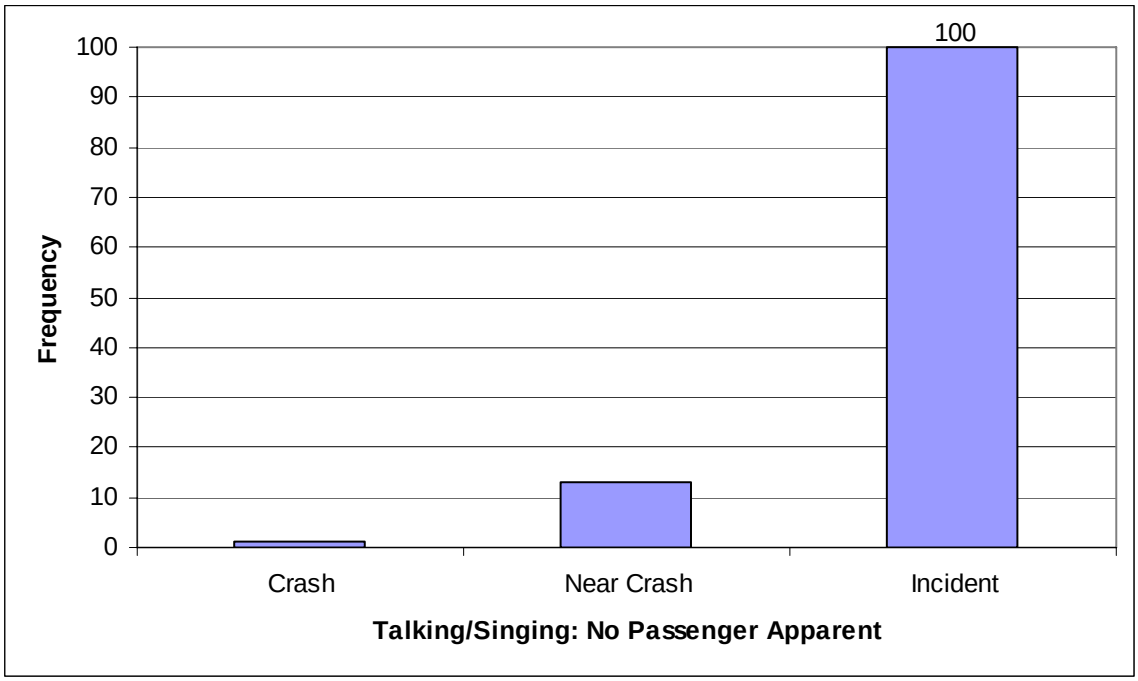


Figure 7.18. Frequency of occurrences in which the contributing factor was talking/singing: no passenger apparent secondary tasks. (Driver N = 241).

External Secondary Tasks. External secondary tasks involved drivers who became interested in something outside the vehicle such as a crash or a construction zone. Appendix D lists descriptions for each of the tasks listed in Figure 7.19. Note that one crash and 10 near-crashes were listed under *other external distraction* meaning that the reductionists could not determine what the driver was observing outside the vehicle. Generally, this category was identified in a relatively low number of events, which contradicts other studies of this type (e.g., Stutts et al., 2003). This may have been partly due to the difficulty in determining whether the driver was observing something specific outside the vehicle or randomly gazing out the window. However, even when considering the external nonspecific eyeglance locations for crash and near-crash events, this category is still lower than other studies have indicated.

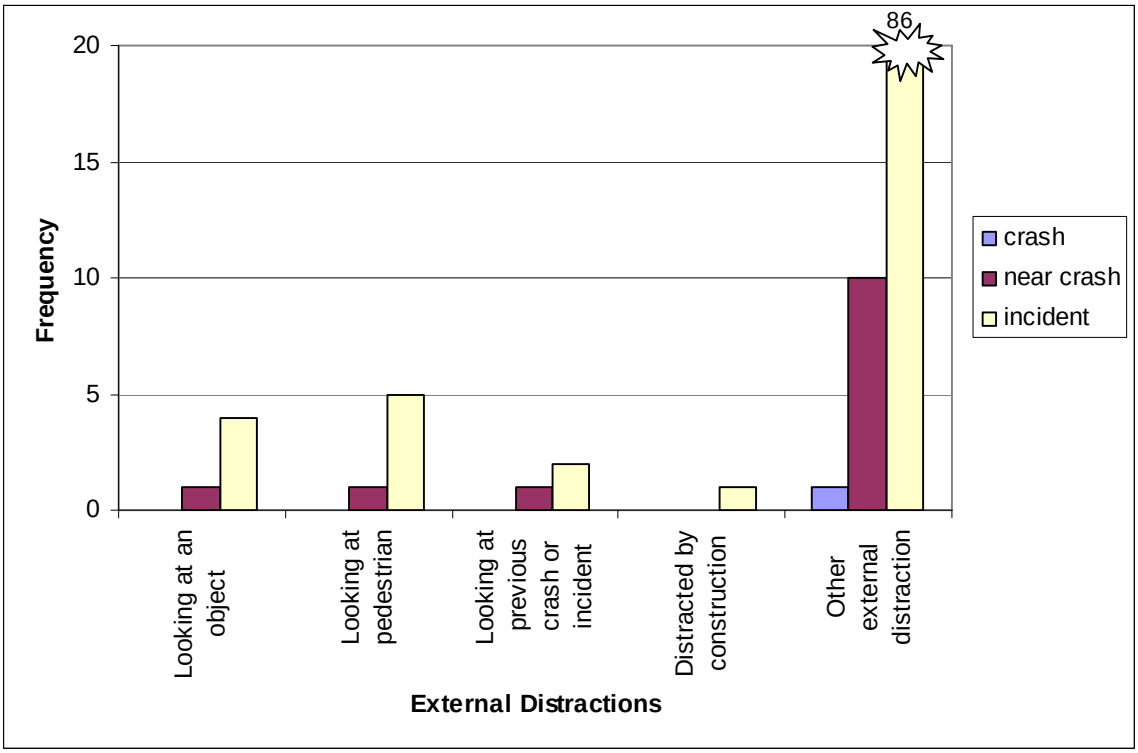


Figure 7.19. Frequency of occurrences in which the contributing factor was an external secondary task (Driver N = 241).

Internal Secondary Tasks: Not Vehicle or Passenger-Related. The internal secondary task category involved the drivers manipulating or locating miscellaneous objects in the vehicle that were not related to in-vehicle systems, other passengers, wireless devices, or specified secondary tasks (such as eating or smoking, which are located under dining and smoking categories). The task types under this category were less frequent compared to other categories (i.e., well under 100 occurrences for each type). All of these categories are defined in Appendix D. Note that both *reaching for an object* having an *object or animal in the vehicle,* and *moving object in vehicle* contributed to crashes. The first two of these categories, plus *reading*, were the primary contributors to near-crashes (Figure 7.20).

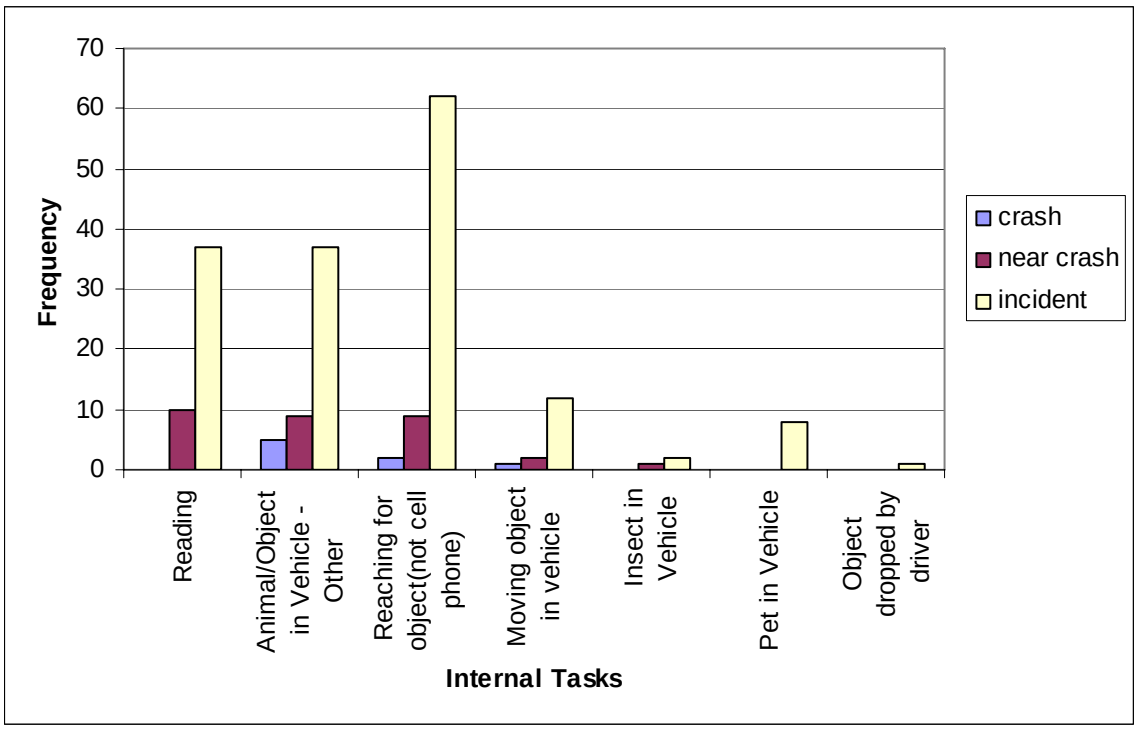


Figure 7.20. The frequency of occurrences in which the contributing factor was internal secondary tasks (Driver N = 241).

Personal Hygiene. This category of secondary tasks pertained to the driver engaging in any grooming, cleaning, or attending to themselves as opposed to being attentive to a miscellaneous object. Again, note that all of the personal hygiene categories are defined in Appendix D. While the overall frequency for this type of secondary task is not as high as other categories, this category is of interest when determining whether drivers are remembering that they are in an instrumented vehicle. No crashes occurred while drivers were engaged in personal hygiene, but a number of near-crashes occurred (Figure 7.21).

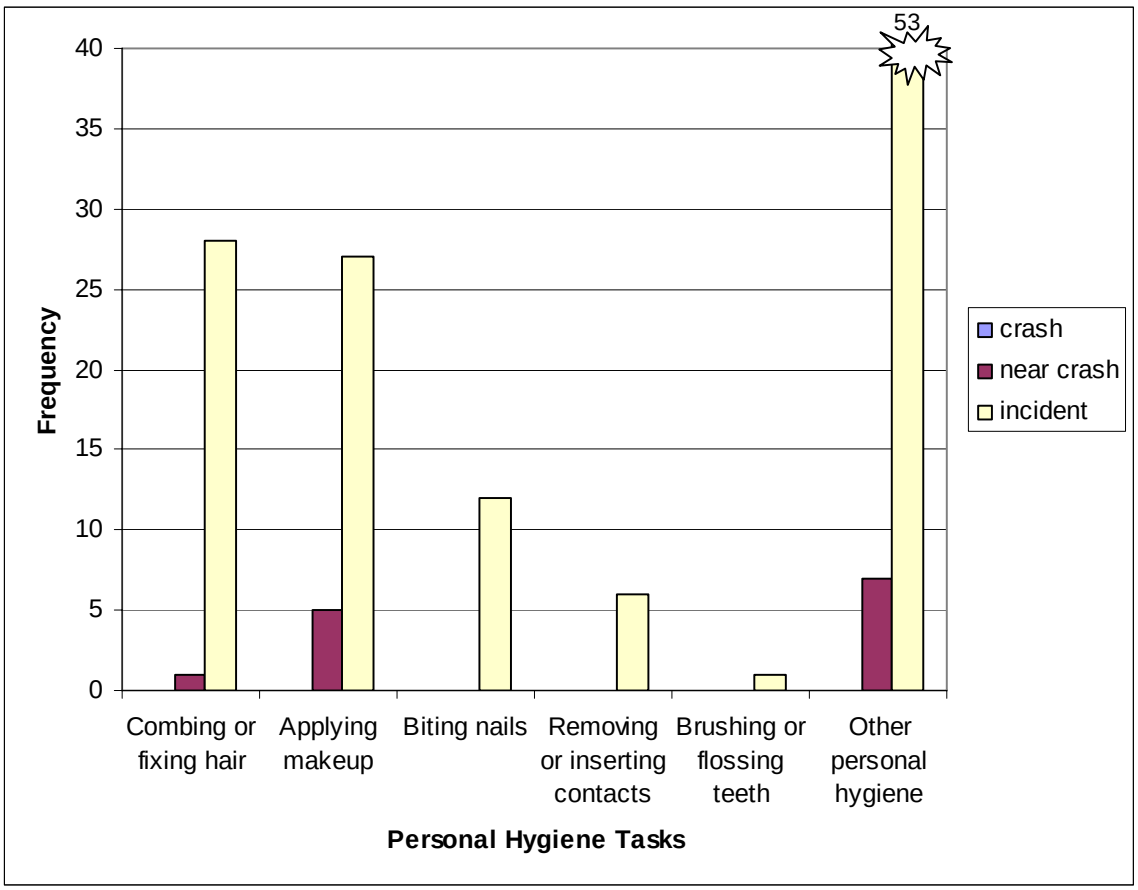


Figure 7.21. Frequency of occurrences in which the contributing factor was a personal hygiene task (Driver N = 241).

Dining. For the secondary task category of *dining*, a distinction was made between drinking out of a covered versus open container (i.e., with or without a lid) and eating with or without utensils. Note that in Figure 7.22, eating without utensils and drinking from an open container both contributed to more crashes and near-crashes than drinking from covered containers or eating with utensils. It is intuitive that drinking from an open container would contribute to more overall events since it is a more difficult task than drinking from a container with a lid. However, it also makes intuitive sense that eating with utensils would be more difficult than eating without utensils, and yet the frequencies show more higher-severity events when eating without utensils. Given that most meals in a vehicle are eaten without utensils, there are probably different levels of exposure for these two tasks.

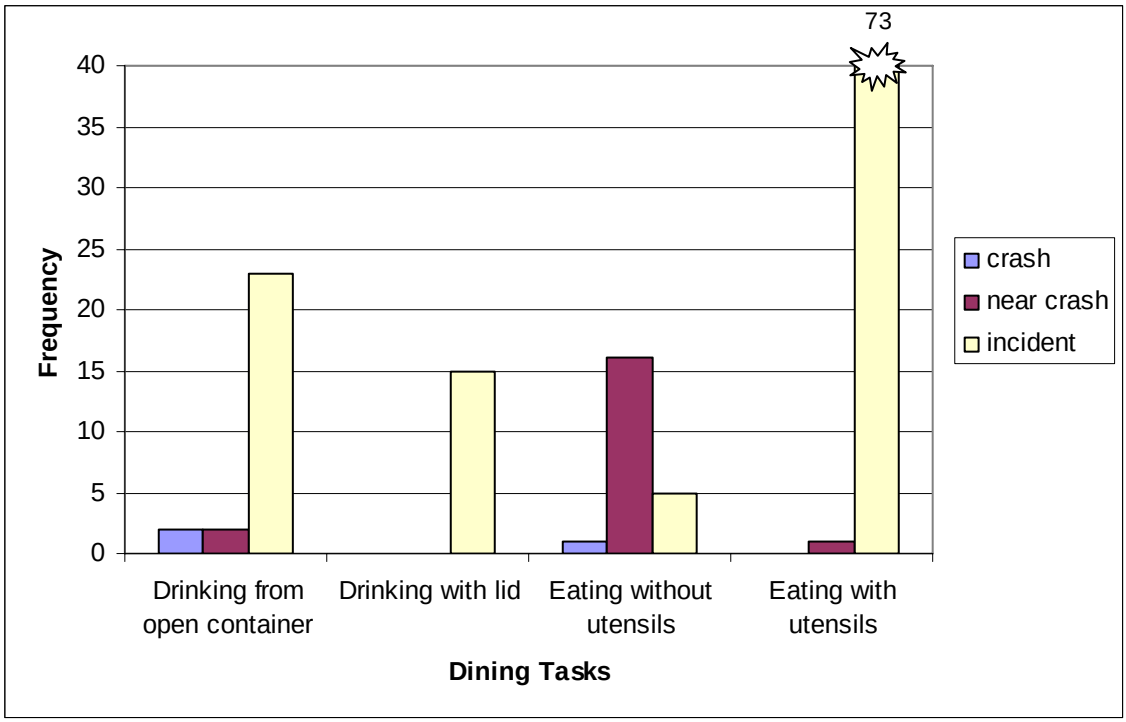


Figure 7.22. Frequency of occurrences in which the contributing factor was a dining task (Driver N = 241).

Smoking. Few occurrences listed smoking as a contributing factor (Figure 7.23). Note that the act of smoking was a contributing factor more often than lighting or reaching for a cigar or cigarette. Incidentally, one rear-end struck collision was caused by a driver in the other vehicle who was lighting a pipe. The driver admitted this to the police officer; however, that event is not recorded here as all events are based upon the instrumented vehicle driver. For the one near-crash that occurred, the driver was smoking a cigar/cigarette and looked away from the forward roadway prior to the onset of the conflict.

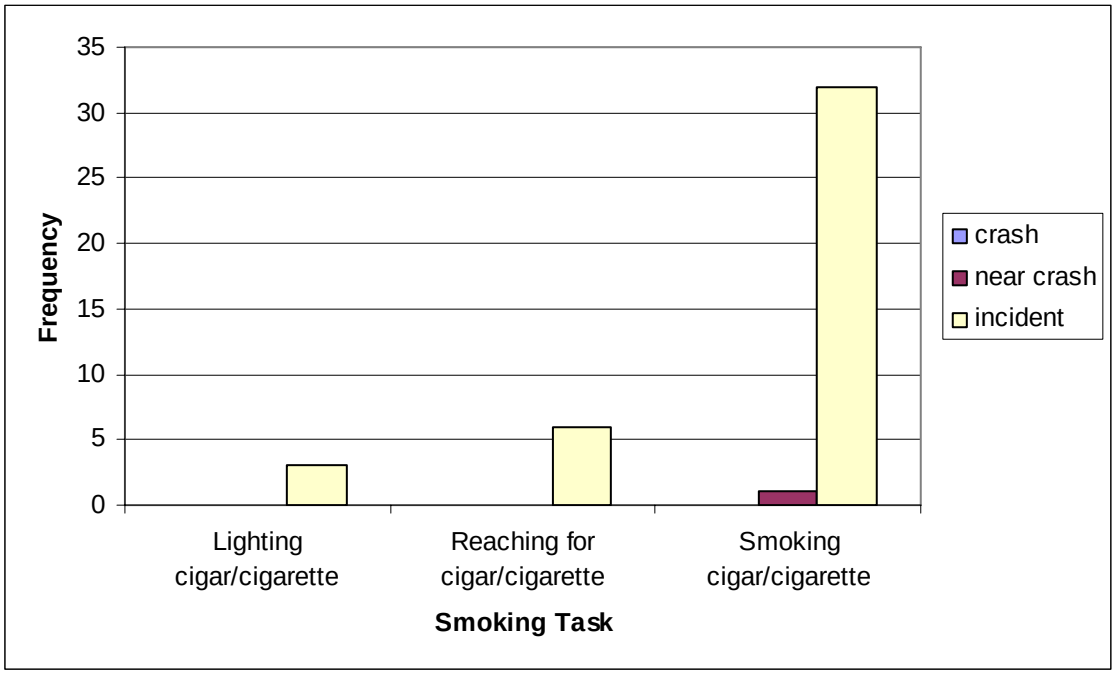


Figure 7.23. Frequency of occurrences in which the contributing factor was a smoking task (Driver N = 241).

Daydreaming. Figure 7.24 shows the number of occurrences for which the reductionists believed the driver was either lost in thought or looked in the direction of but did not observe the conflict (see Appendix D for more detailed definitions of these categories). The low frequency counts reflected the difficulty in assessing whether a driver was daydreaming by simply examining the video. Therefore, the true frequency of daydreaming is probably higher than shown in Figure 7.24.

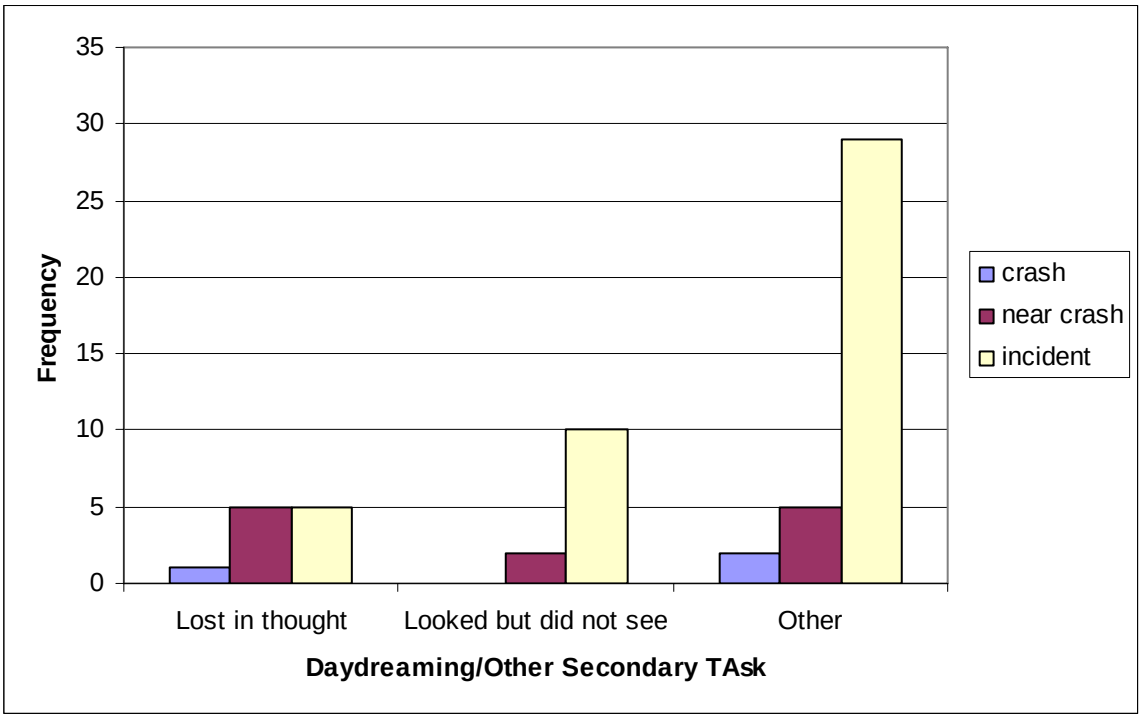


Figure 7.24. Frequency of occurrences in which the contributing factor was a daydreaming task (Driver N = 241).

Question 3. For the incidents, near-crashes, and crashes for which inattention was a contributing factor, what is the prevalence of other driving behaviors occurring such as willful behavior, driver impairment, or drowsiness?

This analysis was conducted to determine if the occurrence of any particular driver (e.g., aggressive driving) behavior tended to increase the level of severity of the event. Also, did the occurrence of inattention plus aggressive driving lead to a near-crash or crash event? For this analysis, percentage values were calculated based upon the total number of events (regardless of attention level).

Data reductionists identified those events for which the driver was demonstrating willful behavior, which was one or a combination of several of the following: aggressive driving, willful violation of traffic laws, or use of vehicle for purposes of intimidation. Driver proficiency was recorded by data reductionists when they observed drivers violating traffic laws or controlling the vehicle in a manner such that it was assumed that the driver lacked knowledge. Examples included consistently driving in an unsafe manner (i.e., stopping or braking suddenly without

cause) and attempting to perform maneuvers for which the vehicle was not designed (i.e., attempting a U-turn without enough available roadway). Driver drowsiness was briefly discussed previously, but is also discussed in this section in conjunction with *driver physical/mental impairment. Driver physical/mental impairment*, similar to GES variable D3, Driver Physical/Mental Condition, was recorded by reductionists to include drowsiness as well as anger, other emotional states, drug or alcohol use, etc. Reductionists only specified drug or alcohol use when it was explicit or when the driver admitted to the behavior during the debriefing process.

Figure 7.25 shows the frequency of willful behavior, driver impairment, and driver proficiency events. The figure shows that driver proficiency appeared to be more problematic than aggressive driving or driver impairment for all levels of severity.

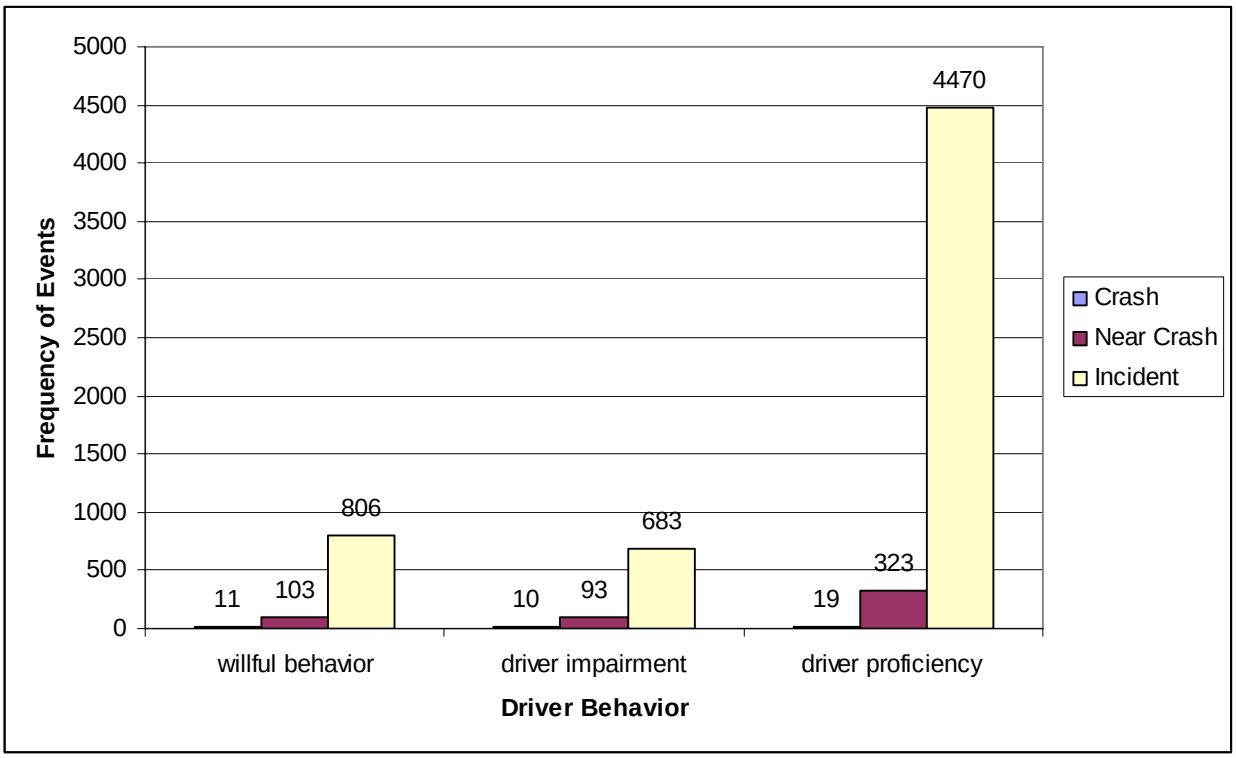


Figure 7.25. Frequency of driver inattention, willful behavior, and driver impairment on the total number of crashes, near-crashes, and incidents in the 100-Car Study database.

Figure 7.26 shows the frequency of the various levels of driver impairment. Please note that drowsiness generally accounted for 22 percent of all events whereas the rest of the impairments were fairly infrequently identified. Without actual interviews for each event, this type of information was difficult to obtain via video reduction.

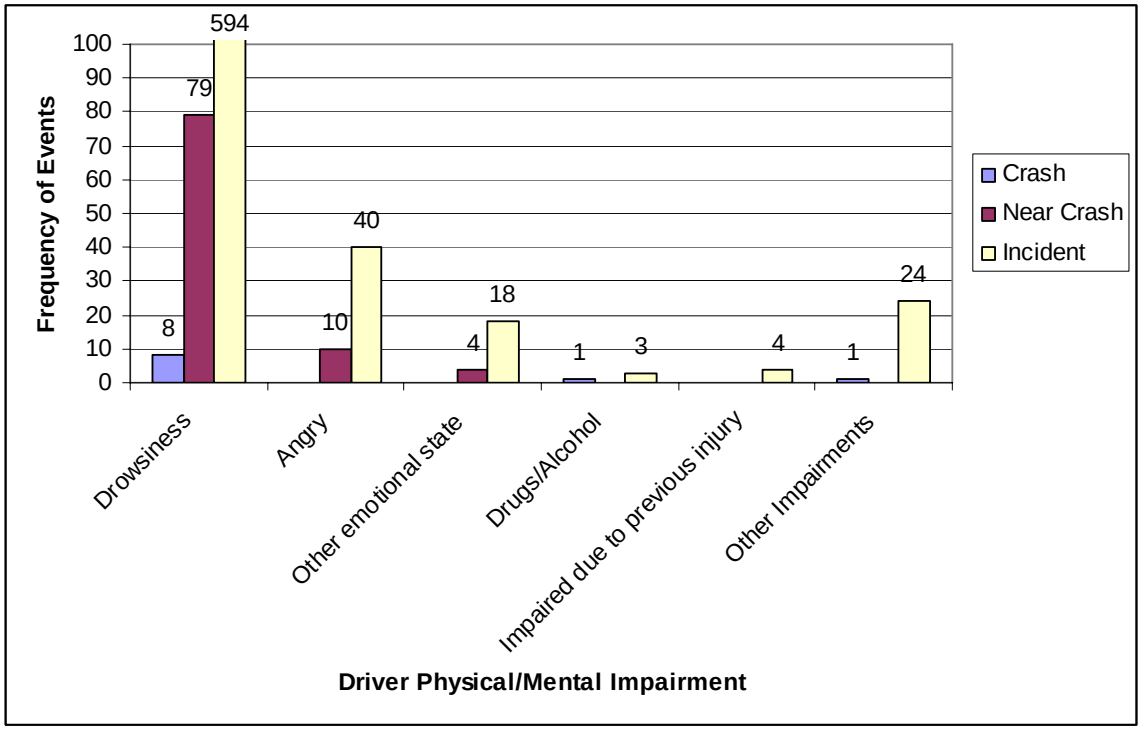


Figure 7.26. Frequency of the driver physical/mental impairment categories listed as a contributing factor.

Figure 7.27 shows the frequency of inattention plus driver behavior for crashes, near-crashes, and incidents. While inattention and driving proficiency were paired together frequently, inattention and willful behavior or driver impairment were less so, although 5 crashes were attributed to the combination of *inattention* and *willful behavior*.

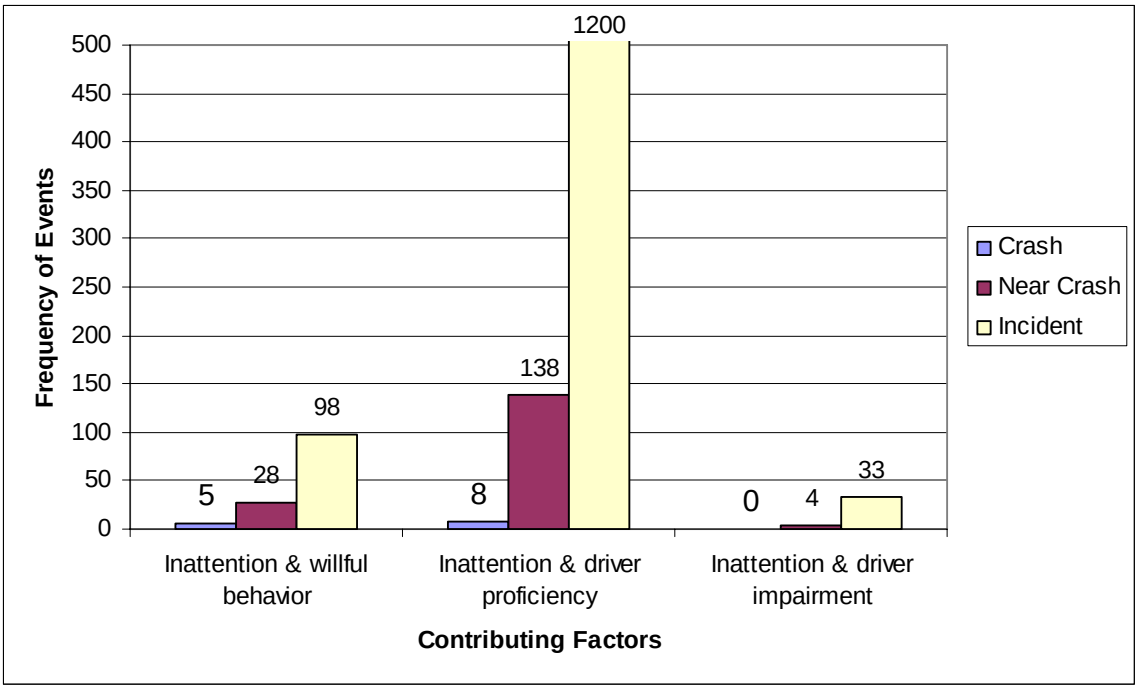


Figure 7.27. The percentage of events that included both inattention and willful behavior, driver proficiency, or driver impairment.

Question 4. Do drivers conversing on cell phones exhibit poorer driving performance than drivers not using cell phones?

While comparisons could potentially be made of drivers on the cell phone to an estimated baseline dataset, it would be more appropriate to address this question using actual baseline driving data. Baseline driving data will be identified and reduced as part of a follow-on effort. Specifically, events will be reduced for which no conflicts occurred and reductionists will record whether the driver is or is not talking on a cell phone.

DISCUSSION

Historically, driver distraction, as well as possibly driver drowsiness, has been typically discussed as a secondary task engagement. In this paper the definition of driver distraction has been expanded to a more encompassing *"driver inattention"* construct by including three new categories, "*driving-related inattention to the forward roadway*," and "*nonspecific eyeglance'*. "*Driver-related inattention to the forward roadway"* involves the driver checking rear-view mirrors or their blind spots. This new category was added after viewing multiple crashes, near-crashes, and incidents for which the driver was clearly paying attention to the driving task, but was not paying attention to the *critical aspect* of the driving task (i.e., forward roadway).

A second analysis of the crashes and near-crashes in the 100-Car Study database was also conducted using the eyeglance analysis performed manually by data reductionists. The *"non specific eyeglance away from forward roadway"* describes cases in which the driver glances, usually momentarily, away from the roadway at a non-discernable object. Cases where the object could be identified were classified as either *driving-related inattention to the forward roadway* or *secondary task distraction* depending on what the driver was looking at. For this project, eyeglance reduction was accomplished for crash and near-crash events, so this category can only be used for the more severe categories. This analysis suggested that driver's glances away from the forward roadway potentially contribute to a much greater percentage of events than the first three categories of inattention suggest.

Thus, with the addition of these two categories, driver inattention has been operationally defined as including drowsiness, engagement in secondary tasks, driving-related inattention to the forward roadway (checking blind spots), and nonspecific eyeglance way from the forward roadway. All of these events were identified by reductionists viewing the driver's behavior surrounding the onset of the event.

It is important to note that the data presented in this section represents raw frequencies or percentages derived from raw frequencies, thus exposure is not accounted for. Specifically, one needs to determine the frequency and duration with which each of these categories of inattention are present during normal, non-event, driving in order to make judgments about relative risk. As of this writing, an additional analysis is underway that will establish exposure and make such relative risk comparisons.

Several very important results were identified as part of this analysis.

- For the crashes and near-crashes, the driver looked away from the forward roadway at least once in a 4 seconds window surrounding the events (3 seconds prior and 1 second post-event onset) in almost 80 percent of the crash cases and 65 percent of the near-crash cases.

- The rate of inattention-related crash and near-crash events decreases dramatically with age, with the rate being as much as four times higher for the 18- to 20-year-old age group relative to some of the older driver groups (i.e., 35 and up).

- The use of hand-held wireless devices (primarily cell phones but including a small presence of PDA use) was associated with the highest frequency of secondary task inattention-related events. This was true for both events of lower severity (i.e., critical incidents) and for events of higher severity (i.e., near-crashes). Wireless devices were also among the categories associated with the highest frequencies of crashes and minor collisions, along with looking/reaching for an object in vehicle and passenger-related secondary tasks.

- *Drowsiness* and *driving-related inattention to forward roadway* were also listed among the most frequent contributors to crashes and near-crashes. The driving-related

inattention to the forward roadway category: *looking out the left window* was associated with the highest number of events for this classification. Both of these categories were higher than expected based upon previous research or conventional wisdom.

Driver inattention did not appear to combine meaningfully with driver mental/physical impairment, willful behavior, or driver proficiency to contribute to events, although these behaviors on their own did contribute to many of the events in the 100-Car Study database.

A few issues should be noted when interpreting the above results. Secondary tasks, drowsiness, and inattention to forward roadway were recorded by reductionists as objectively as possible. For example, *passenger in vehicle* was recorded when a reductionist observed the presence of a passenger in conjunction with the driver reacting either late or in an inappropriate manner. Therefore, the fact that an event occurred with a passenger in the vehicle cannot be interpreted as a *cause* of the event; instead, the presence of a passenger can only *correlated* with an event. This is true for every secondary task, including the use of a cell phone. These correlations will be further analyzed in follow-on work. Driver performance during inattention and baseline events will be used to perform statistical tests for which inferences can then be discussed.

Eyeglance analysis, while manually performed by reductionists, is an objective task that is somewhat less prone to human error. Nevertheless, a driver's glance direction, even if focused on the onset of an event, provides no guarantee that the driver will see or perceive the situation appropriately. While human error cannot be entirely removed from this type of data collection, reductionists determination of inattention and eyeglance data are both required to shed light on the problem of driving inattention.

CHAPTER 8: GOAL 4, DRIVER PERFORMANCE IN INSTRUMENTED VEHICLES OVER TIME: OVER THE FIRST FEW HOURS, OVER THE FIRST YEAR, AND OVER ONE MONTH FOR SAME DRIVER IN PRIVATE VERSUS LEASED VEHICLES

DATA ANALYSIS OVERVIEW

The questions addressed in Chapter 8, *Goal* 4 were intended to explore issues of whether driver behavior in an instrumented vehicle changed over time. The units of time used were weeks (weeks 1 through 50) and hours (the first 50 hours). The issues explored were driver behavior in a newly instrumented leased vehicle in the first weeks as compared to driver behavior in the first few hours of driving, and driver behavior for the same driver in four weeks of leased vehicle driving and four weeks of private vehicle driving by the end of the study.

The analyses discussed in this chapter were conducted using epidemiological methods. The definitions and formulas presented are from Greenberg et al. (1993). Some of the terms and assumptions have been modified to fit the current dataset and analyses. The 100-Car Study was deemed to most closely fit the definition of a cohort study, in which investigators identify an initial population and determined their initial exposure status. Groups were then exposed to different conditions and tracked over time. So in the case of the 100-Car Study, the initial cohort was the group of drivers who agreed to participate. If the exposure of interest was exposure to a leased vehicle (as opposed to the private vehicles they had all been driving up to this time), then some of the drivers would be exposed to a leased vehicle while others would remain in their private vehicles. Driving behavior would then be tracked over time.

According to Greenberg et al. (1993), the appropriate analysis for this type of study is the risk ratio (sometimes called relative risk; it will be referred to as RR in this report). This requires the data to be put into the form shown in Table 8.1.

Table 8.1. Example data matrix for calculating risk ratio. A, B, C, and D are the numbers of cases satisfying each criteria.

Outcome	**Exposed**	**Unexposed**	**Total**
Week of Interest	A	B	A + B
Control Week	C	D	C + D
Total	A + C	B + D	A + B + C + D

The formula for calculating RR is:

$$RR = \frac{R_{(\text{exposed})}}{R_{(\text{unexposed})}} \qquad \text{Eq. 1}$$

in which $R_{(\text{exposed})}$ = *the risk of an outcome for an exposed person* and $R_{(\text{unexposed})}$ = *the risk of an outcome for an unexposed person*. The formula for calculating these risks takes the form:

$$R_{(\text{exposed})} = \frac{A}{A+C} \qquad \text{Eq. 2}$$

in which A and C are obtained from Table 8.1. The formula for calculating RR can thus be expressed as:

$$RR = \frac{A/(A+C)}{B/(B+D)} \qquad \text{Eq. 3}$$

For each RR calculated in this way, the 95th percentage confidence intervals can then be calculated using the following formula:

$$95\%CI = \exp \pm 1.96\sqrt{\frac{1-A/(A+C)}{A} + \frac{1-B/(B+D)}{B}} \qquad \text{Eq. 4}$$

These confidence intervals tend to be asymmetric for RR calculations (with a lower bound of 0 and no upper bound).

For the 100-Car Study dataset, one of the primary questions in calculating RR was the decision regarding which period(s) of time to use as a control period for each question of interest. A preliminary examination showed that the data were rather noisy over time, so a decision was made to use an average of the final 10 time periods for each dataset as the control period. For the analyses performed by week, weeks 41-50 were used, while for the analyses by hour, hours 41-50 were used.

With regard to the research questions for Chapter 8, *Goal* 4, the first set of three questions attempts to characterize these differences based on raw numbers (actual numbers of valid events, without regard to exposure), while the second set of three questions does the same on a per mile basis (attempting to control for exposure).

Data Included in the Analyses

The analyses for Chapter 8, *Goal* 4 included incidents recorded for all primary drivers of vehicles. There were 109 primary vehicle drivers, but due to data outages, the number of drivers available for any given week ranged from a high of 107 in week 1 to a low of 60 in week 49. There were 64 drivers in week 50. Altogether, there were 8,229 events available for the weekly analyses (weeks 1 through 50). Of these, 7,472 were incidents, 696 were near-crashes, and 61

were crashes. For the hourly analyses, there were 974 events available in hours 1 through 50. Of these, 845 were incidents, 118 were near-crashes, and 11 were crashes.

Relevant variables used during the analyses included: age (younger and older, in which younger is 30 or younger and older is older than 30); leased vehicle versus private vehicle, switch driver versus non-switch driver (in which a switch driver is a private vehicle driver who was given a leased vehicle for four to eight weeks at the end of the study); and event severity (crashes, near-crashes, and incidents, defined as usual in this report).

Question 1. Based on the number and type of valid events, is there a significant difference in the relative risk of driving over the course of a year for drivers in a familiar vehicle with instrumentation installed (leased and private vehicles)?

The purpose of this question is to investigate the driver adaptation process for leased vehicles and privately owned vehicles with instrumentation over the course of the study. It was desired that the control time period be near the end of the study, when drivers would have become most adapted to their vehicles. Thus, relative risk was calculated using an average of weeks 41-50 as the control time period.

Figure 8.1 shows the mean number of incidents for private versus leased vehicles for the entire time period of interest in this analysis (weeks 1-50). It was originally proposed that weeks 1, 4, 12, and 26 be used as the weeks of interest for the RR analysis for this research question. However, as shown in Figure 8.1, the data is quite noisy and the patterns shown by such an analysis will vary depend on which weeks are chosen. For example, Figure 8.2 shows the RR for weeks 1, 4, 12, 26, and 49, with week 50 as the control week. Figure 8.2 can be compared to Figures 8.3 and 8.4 in which different comparison weeks were chosen. The patterns are quite different in the three cases. Therefore, some smoothing of the data was desired for this analysis. The following time periods were thus selected for subsequent analysis and graphing: Week 1; Week 2; Week 3; Week 4; Weeks 1-4; Weeks 5-8; Weeks 9-12; Weeks 13-16; Weeks 17-20; Weeks 21-24; Weeks 25-28; Weeks 29-32; Weeks 33-36; Weeks 37-40; Weeks 41-44; Weeks 45-48; and baseline weeks 41-50. Weeks 1 through 4 were examined individually, as well as averaged, so that the effect of vehicle adaptation to a strange vehicle would not be masked, if such an effect were indeed present.

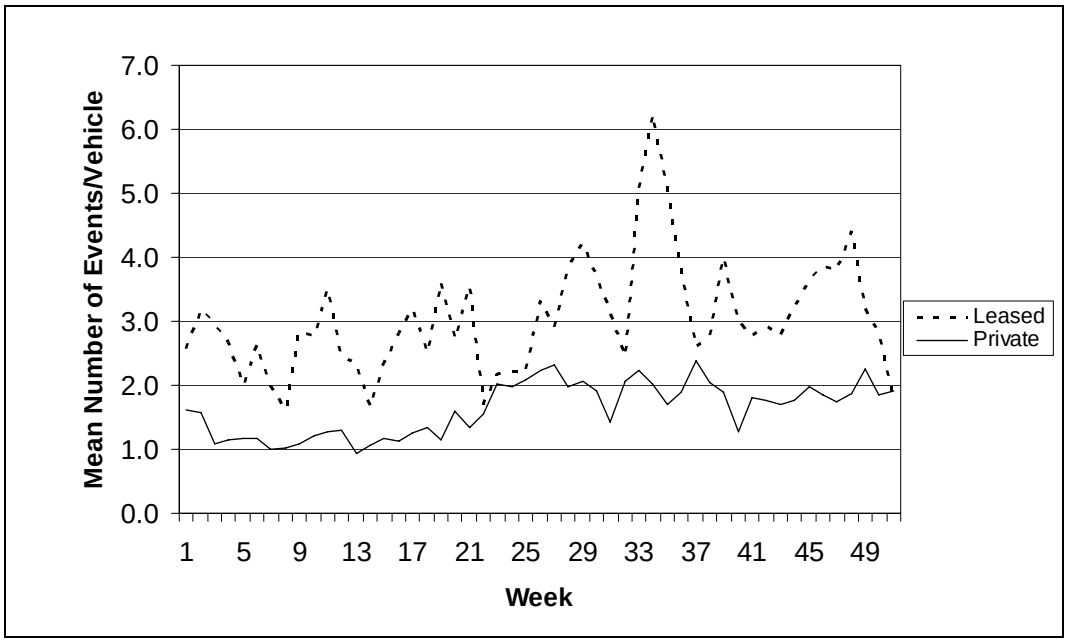


Figure 8.1. Mean number of events per vehicle for weeks 1-50 for leased and private vehicles.

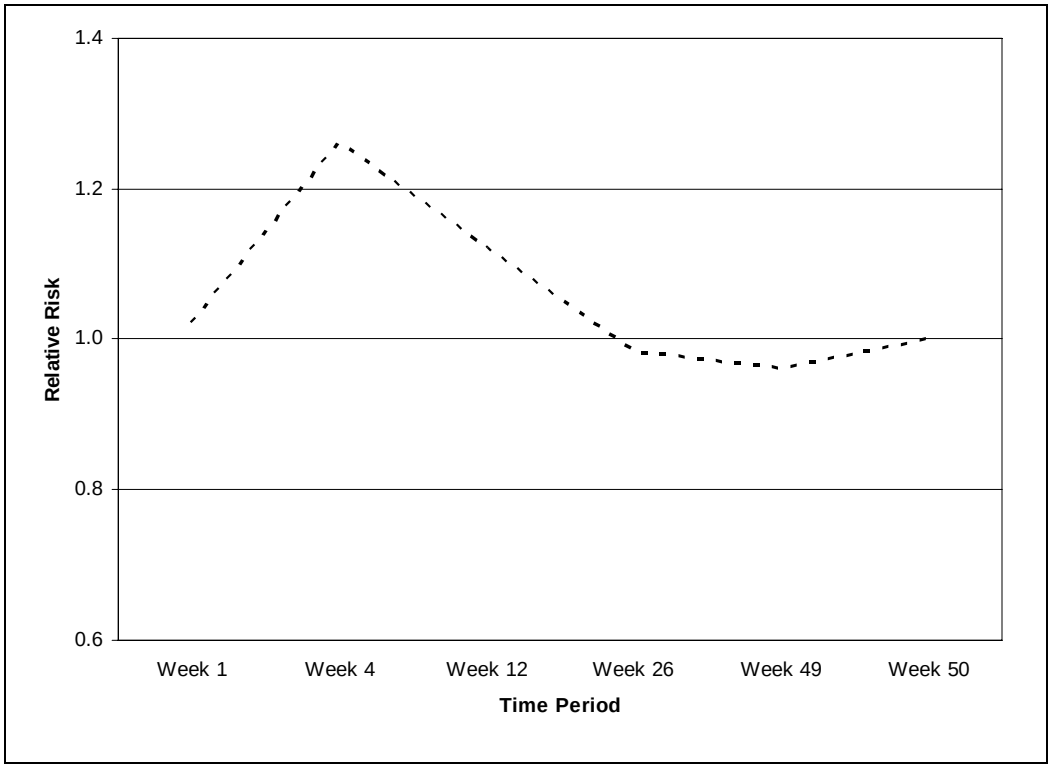


Figure 8.2. Comparison of RR using the originally proposed comparison weeks of 1, 4, 12, 26, and 49 using week 50 as a control week.

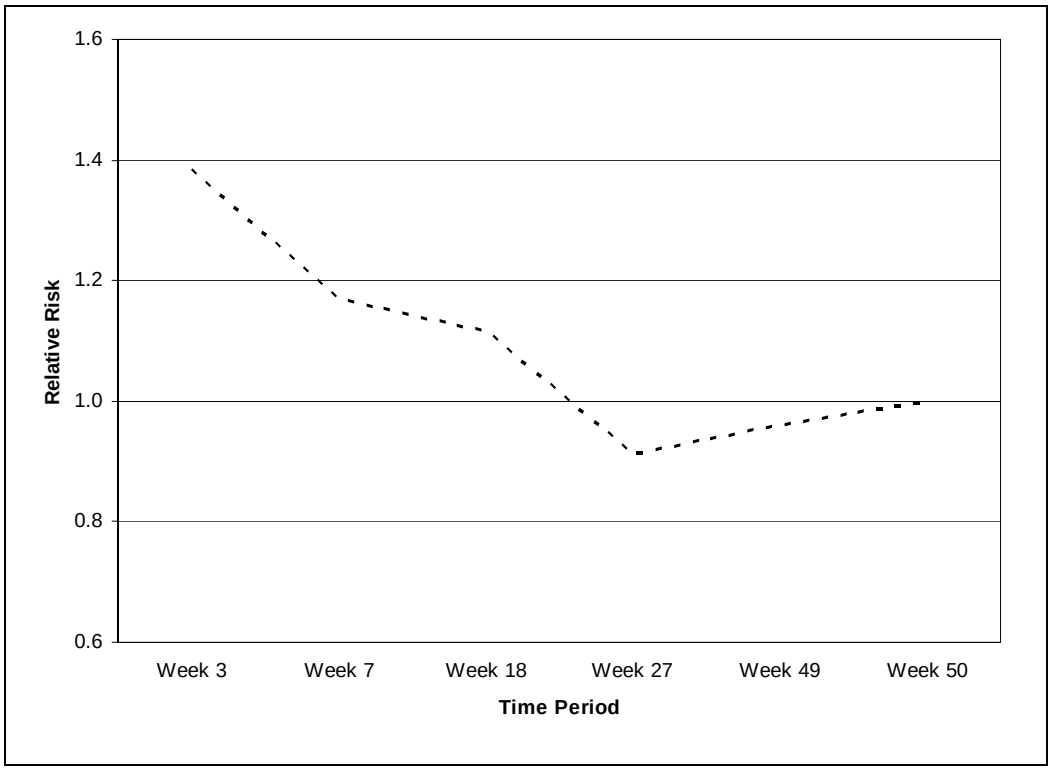


Figure 8.3. Comparison of RR using weeks of 3, 7, 18, 27, and 49 using week 50 as a control week and showing a strong downward trend.

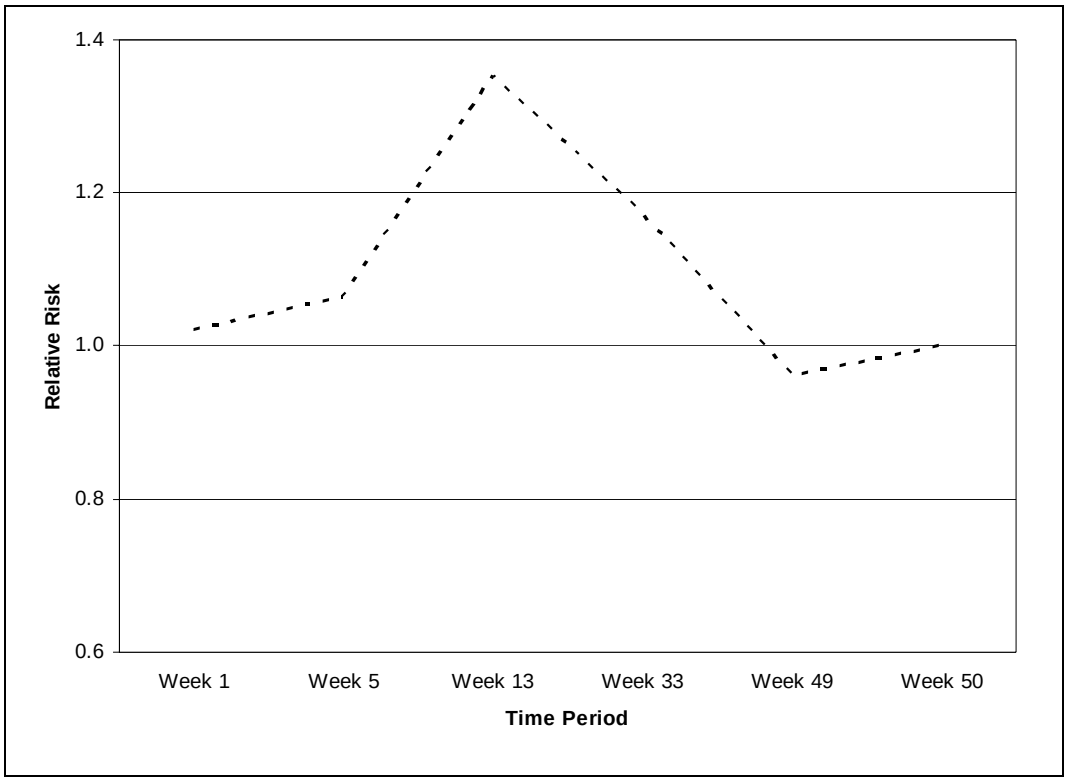


Figure 8.4. Comparison of RR using weeks of 1, 5, 13, 33, and 49 using week 50 as a control time period and showing an upward trend followed by a downward trend.

An examination of Figure 8.1 above shows some interesting features. For example, up to around week 23, the leased vehicles had about twice the mean rate as the private vehicles. Upon first glance, one might be tempted to interpret this as some sort of seasonal variation. However, it should be noted that the week denoted in the dataset referred to that driver/vehicle combination's week of driving since entering the study. This eliminated seasonal variations. For example, if the Washington, DC, area experienced a snowstorm on January 12, and there was an unusually high number of incidents noted during that time frame, this should not appear in the graphs organized by week. That is, January 12 may have been part of week 1 for driver 45, week 17 for driver 32, and week 26 for driver 65. This method should also have eliminated any other possible systematic variations that could have occurred over time.

At about week 23, the private vehicles began to experience a higher mean rate, which continued through the remainder of the study. The leased vehicles also began experiencing higher mean rates at about that time, but the increase was not quite of the same magnitude as for the private vehicles. The most interesting thing about Figure 8.1, however, is the consistency of the difference between the leased vehicle mean rate and the private vehicle mean rate. This pattern would indicate that leased vehicles would have a high risk ratio when they are compared to private vehicles.

Table 8.2 presents the mean number of events for private and leased vehicles for the time periods of interest as described previously. The mean number of events was calculated by dividing the number of events for time period of interest by the number of drivers for same time period so

that the lower number of leased vehicles as compared to private vehicles is accounted for in the calculation.

Table 8.2. Mean number of events per vehicle per time period for leased and private vehicles; calculated relative risk for those time periods; 95th percentage confidence intervals based on RR.

Time Period	Leased Average	Private Average	Relative Risk
Week 1	2.57	1.61	0.93
Week 2	3.19	1.57	1.06
Week 3	2.95	1.08	1.27
Week 4	2.66	1.15	1.15
Weeks 1-4	2.83	1.36	1.08
Weeks 5-8	2.07	1.09	1.03
Weeks 9-12	2.90	1.22	1.17
Weeks 13-16	2.29	1.07	1.11
Weeks 17-20	3.01	1.33	1.13
Weeks 21-24	2.42	1.72	0.87
Weeks 25-28	3.09	2.16	0.89
Weeks 29-32	3.35	1.86	1.00
Weeks 33-36	5.09	1.96	1.17
Weeks 37-40	3.10	1.90	0.95
Weeks 41-44	2.93	1.76	0.96
Weeks 45-48	3.93	1.87	1.07
Baseline Weeks 41-50	3.37	1.86	1.00

In examining Table 8.2, note that there is a large discrepancy between the mean number of events for the baseline leased (3.37) and private (1.86) vehicles. When the RR is calculated using quite different numbers for the baseline (control) part of the equation, the RR is artificially lower than it should be, based on the individual differences for each time period. The idea behind using the relative risk technique in epidemiological studies is that you would not expect to find differences between the groups of interest for the control time period. Based on this principle, we might expect that if we were to transfer the leased vehicle drivers into their own private vehicles in which they would be responsible for repairs, insurance, etc., that their incident levels would drop to the same levels shown for private vehicle drivers. For the remainder of these analyses, the baseline time period for the relative risk calculation was thus set equal to the value for the private vehicles during the same time period. This had the effect of raising the RR to more realistic levels based on the graphs of mean events. Table 8.3 presents the revised mean number of events for private and leased vehicles for the time periods of interest as described previously. Note that the RR is now greater than 1 for every time period, in which it was less than 1 for about half of the time periods in the original Table 8.2. The 95th percentage upper and lower confidence intervals for weeks 1-50 are presented at the end of the table. The mean number of events for private and leased vehicles for these time periods is presented in Figure 8.5, while the RR for these time periods is shown in Figure 8.6. The 95th percentage confidence intervals are overlaid on top of the RR graph in Figure 8.7. Note that when placed in the context of 95th percentage confidence intervals as in Figure 8.7, the variation in the RR itself is shown to be quite small, although larger than it would be if calculated with the initial baseline time period.

Table 8.3. Mean number of events per vehicle per time period for leased and private vehicles; calculated relative risk for that time period. The baseline leased vehicle mean is now set equal to the private vehicle baseline.

Time Period	Leased Average	Private Average	Relative Risk
Week 1	2.57	1.61	1.25
Week 2	3.19	1.57	1.38
Week 3	2.95	1.08	1.67
Week 4	2.66	1.15	1.54
Weeks 1-4	2.83	1.36	1.43
Weeks 5-8	2.07	1.09	1.42
Weeks 9-12	2.90	1.22	1.54
Weeks 13-16	2.29	1.07	1.51
Weeks 17-20	3.01	1.33	1.48
Weeks 21-24	2.42	1.72	1.18
Weeks 25-28	3.09	2.16	1.16
Weeks 29-32	3.35	1.86	1.29
Weeks 33-36	5.09	1.96	1.43
Weeks 37-40	3.10	1.90	1.23
Weeks 41-44	2.93	1.76	1.26
Weeks 45-48	3.93	1.87	1.35
Baseline Weeks 141-50	1.86	1.86	1.00
Weeks 1-50	2.95	1.60	1.33
Weeks 1-50 95% CI: Upper, Lower		5.06	0.35

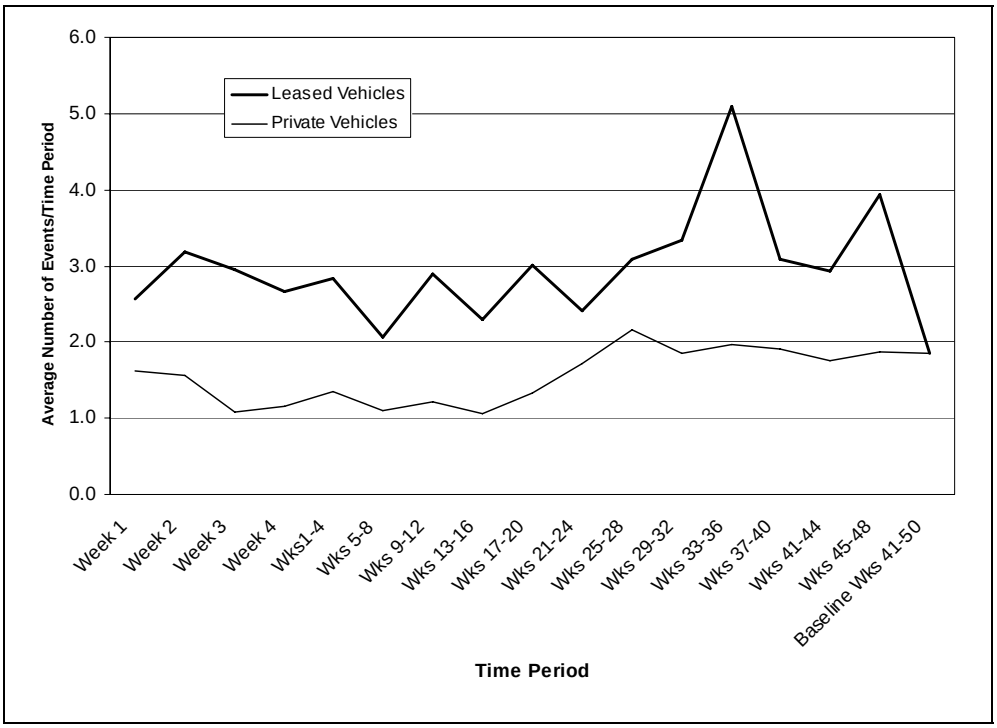


Figure 8.5. Mean number of events per vehicle for leased and private vehicles for time periods of interest over weeks 1-50 of the study.

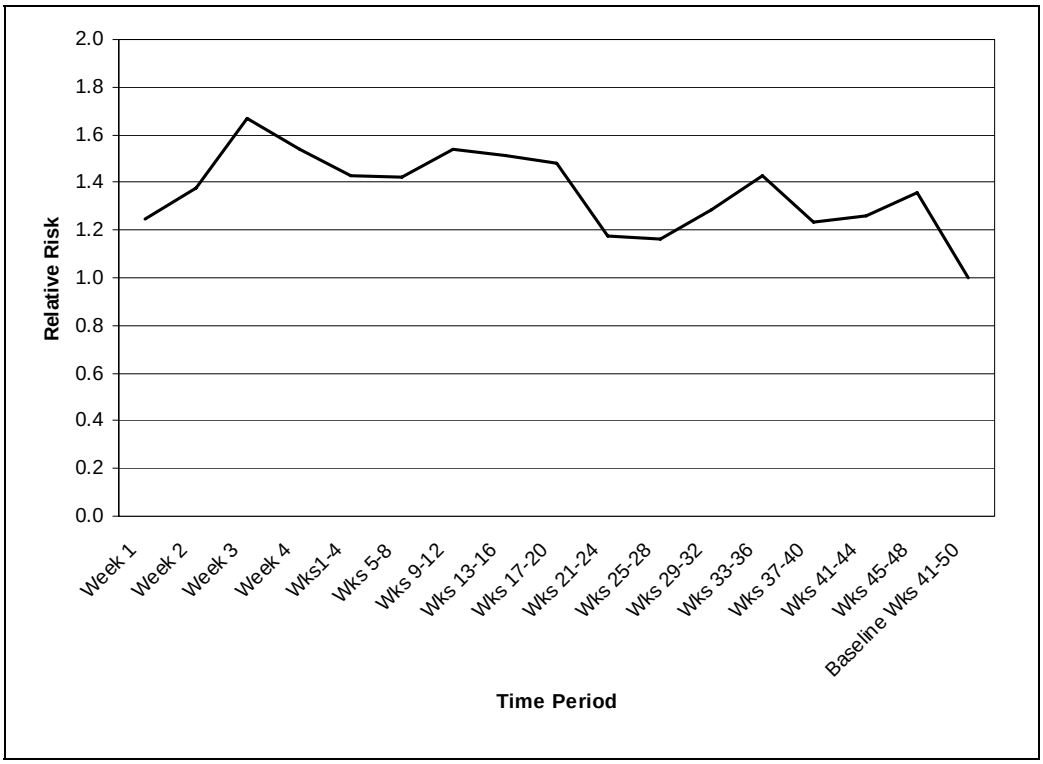


Figure 8.6. Relative risk for leased versus private vehicles for time periods of interest over weeks 1-50 of the study.

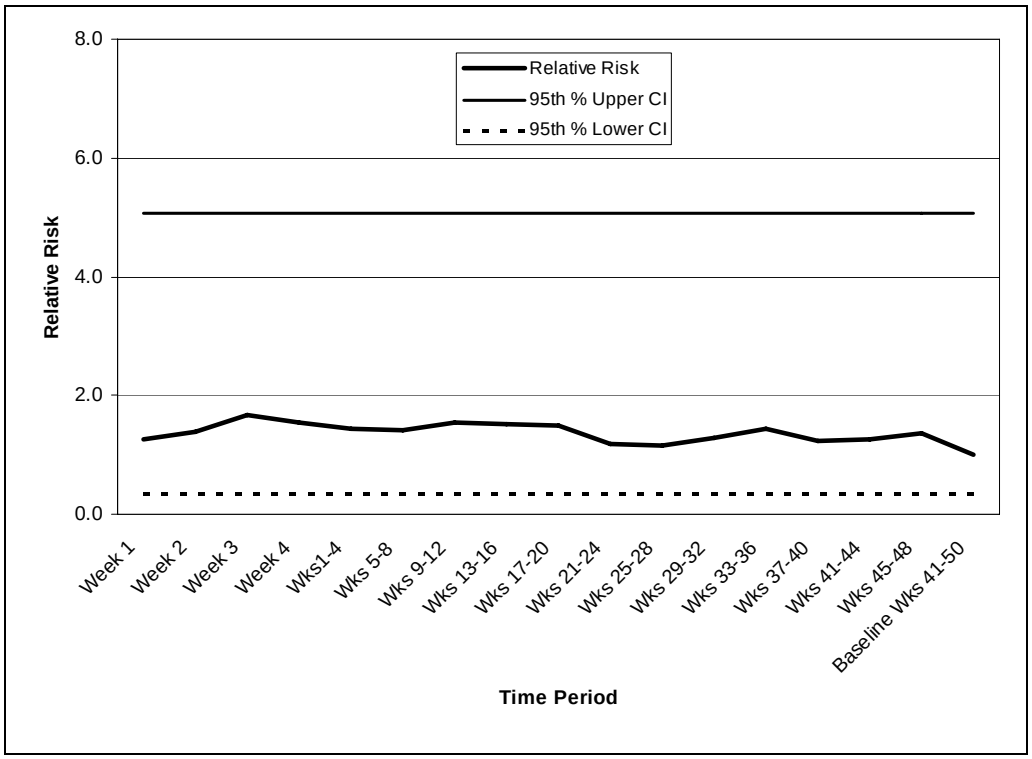


Figure 8.7. Relative risk and 95th percentage confidence intervals for leased versus private vehicles for time periods of interest over weeks 1-50 of the study.

It should be noted that one possible explanation for the higher rates of events for leased vehicles was that there was a large age discrepancy between the two groups of drivers. As discussed in the introductory chapters, none of the leased vehicle drivers were over the age of 30. The mean age of the leased vehicle drivers was 22.3 years (SD = 3.2), while it was 40.7 years (SD = 13.5) for the private vehicle drivers. The age difference may be the most logical explanation for the observed differences in event rate between leased and private vehicles.

Given the possible age-related explanation, the RRs were calculated to all of the older drivers (here considered to be those over age 30, since none of the leased vehicle drivers was over 30) and younger drivers (30 and under), again using weeks 41-50 as the control time period. The data for this age analysis is shown in Table 8.4, including the RR. Figure 8.8 presents the mean number of incidents for younger and older drivers, Figure 8.9 shows the RR for each time period, and Figure 8.10 provides the confidence interval overlay for the RR.

Table 8.4. Mean number of events per vehicle per time period for leased and private vehicles; calculated relative risk for those time periods; 95th percentage confidence intervals based on RR.

Time Period	Younger Average	Older Average	Relative Risk
Week 1	2.63	1.82	1.19
Week 2	2.53	2.17	1.07
Week 3	1.94	1.81	1.04
Week 4	1.98	1.69	1.08
Weeks 1-4	2.28	1.87	1.10
Weeks 5-8	1.63	1.47	1.06
Weeks 9-12	2.36	1.27	1.39
Weeks 13-16	1.79	1.13	1.31
Weeks 17-20	2.38	1.30	1.38
Weeks 21-24	2.35	1.49	1.26
Weeks 25-28	2.72	2.11	1.12
Weeks 29-32	2.38	2.06	1.07
Weeks 33-36	3.72	1.89	1.33
Weeks 37-40	2.58	1.88	1.16
Weeks 41-44	2.16	1.92	1.06
Weeks 45-48	2.66	2.08	1.12
Baseline Weeks 141-50	1.97	1.97	1.00
Weeks 1-50	2.38	1.69	1.18
Weeks 1-50 95% CI: Upper, Lower		4.67	0.30

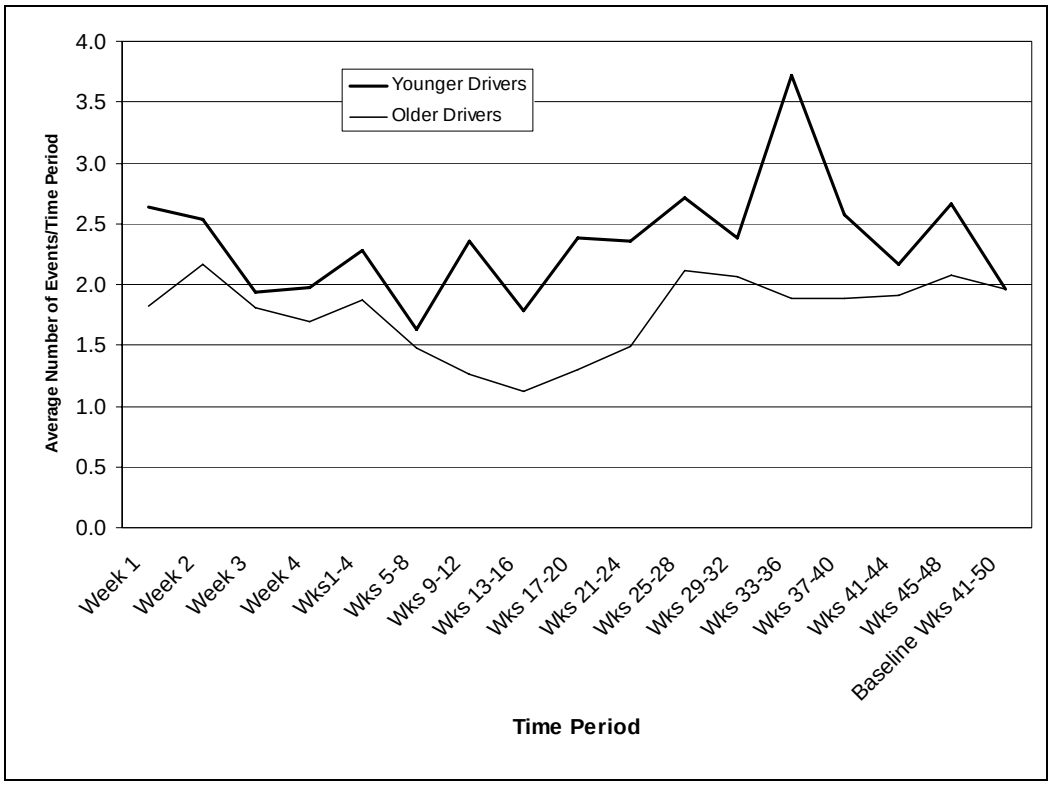


Figure 8.8. Mean number of events per vehicle for younger and older drivers for time periods of interest over weeks 1-50 of the study.

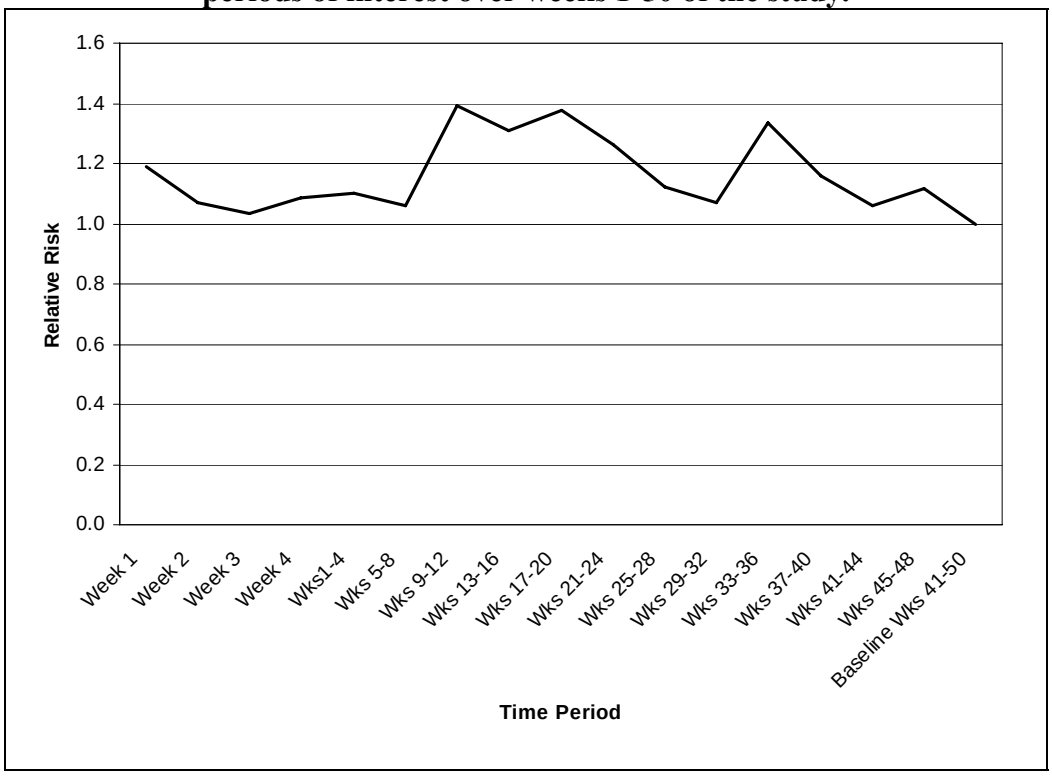


Figure 8.9. Relative risk for younger versus older drivers for time periods of interest over weeks 1-50 of the study.

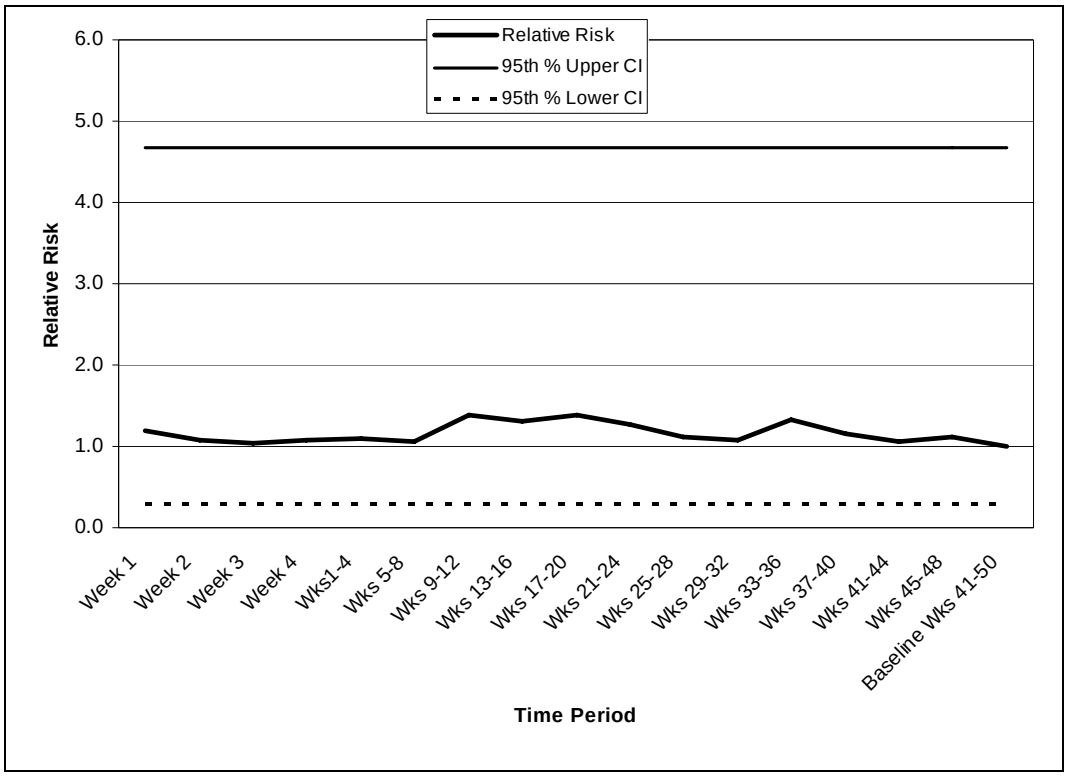


Figure 8.10. Relative risk and 95th percentile confidence intervals for leased versus private vehicles for time periods of interest over weeks 1-50 of the study.

When Figures 8.5 and 8.8 are compared, the patterns for mean number of events or leased versus private vehicles and younger versus older drivers is seen to be quite similar, with younger drivers and leased vehicle having a typically higher average number of events for every time period. However, the magnitude of the difference is larger for leased versus private vehicles than for older versus younger drivers.

The next analysis attempted to eliminate the age factor from the leased versus private vehicle analysis. Since none of the leased vehicle drivers was over 30 (except for switch drivers, but this will be conducted as a separate analysis within Chapter 8, *Goal* 4), the data were examined to see whether there were enough younger drivers to conduct a valid leased versus private vehicle comparison with only the younger drivers. Table 8.5 shows the number of young drivers available in each of these categories for weeks 1-50. Table 8.6 shows the number of drivers for each age from 18-30 for the younger drivers. Taken together, these tables show that the younger drivers were fairly evenly divided in terms of leased versus private vehicles and that the ages were fairly evenly distributed among the leased and private groups. Altogether there were 27 younger leased vehicle drivers and 25 private vehicle drivers. Switch drivers were only counted in the private vehicle column (i.e., they were not double counted when they switched to a leased vehicle at the end of the study).

Table 8.5. Number of leased and private vehicles available for weeks 1-50 for younger drivers only.

Week	Leased Vehicles	Private Vehicles	Total	Week	Leased Vehicles	Private Vehicles	Total
1	27	25	52	26	18	22	40
2	25	24	49	27	20	21	41
3	25	24	49	28	19	21	40
4	26	22	48	29	17	21	38
5	24	22	46	30	16	22	38
6	25	22	47	31	16	22	38
7	27	22	49	32	19	20	39
8	25	21	46	33	17	16	33
9	26	20	46	34	16	16	32
10	26	19	45	35	15	16	31
11	26	19	45	36	13	17	30
12	26	21	47	37	14	17	31
13	25	22	47	38	16	18	34
14	24	21	45	39	15	17	32
15	24	20	44	40	14	17	31
16	25	21	46	41	12	19	31
17	24	22	46	42	13	19	32
18	23	21	44	43	13	17	30
19	22	21	43	44	13	16	29
20	21	20	41	45	13	17	30
21	19	20	39	46	14	17	31
22	19	22	41	47	13	17	30
23	17	21	38	48	14	18	32
24	17	23	40	49	10	18	28
25	17	22	39	50	11	19	30

Table 8.6. Number of younger drivers of each age for leased and private vehicles.

Age	Leased	Private	Total
18	2	1	3
19	5	2	7
20	3	3	6
21	1	0	1
22	2	1	3
23	6	6	12
24	3	2	5
25	1	0	1
26	2	0	2
27	0	3	3
28	0	1	1
29	1	6	7
30	1	0	1
Total Drivers	**27**	**25**	**52**
Mean Age	**22.30**	**24.28**	**23.25**
St. Dev. of Age	**3.16**	**3.74**	**3.56**

Given that the numbers and ages were a reasonable matched set, a comparison of leased versus private vehicles was performed including only the younger drivers. Table 8.7 shows the means and RR for each time period of interest. Figure 8.11 shows the mean number of events for each time period, while Figure 8.12 shows the RR and Figure 8.13 shows the RR with the 95th percentage upper and lower confidence intervals overlaid.

Table 8.7. Mean number of events for younger drivers only per time period for leased and private vehicles; calculated RR for those time periods. Leased vehicle values do not include leased portion of switch drivers.

Time Period	Leased Average	Private Average	Relative Risk
Week 1	2.67	2.12	1.10
Week 2	2.80	1.46	1.34
Week 3	2.32	0.96	1.59
Week 4	1.92	1.05	1.39
Weeks 1-4	2.43	1.41	1.29
Weeks 5-8	2.18	0.97	1.54
Weeks 9-12	3.10	1.18	1.57
Weeks 13-16	2.49	0.92	1.68
Weeks 17-20	3.39	1.37	1.48
Weeks 21-24	2.56	2.19	1.07
Weeks 25-28	3.26	2.26	1.15
Weeks 29-32	3.54	1.45	1.46
Weeks 33-36	5.43	2.12	1.36
Weeks 37-40	3.31	1.96	1.23
Weeks 41-44	3.16	1.45	1.40
Weeks 45-48	4.22	1.43	1.55
Baseline Weeks 141-50	1.65	1.65	1.00
Weeks 1-50	3.14	1.58	1.34
Weeks 1-50 95% CI: Upper, Lower		4.87	0.37

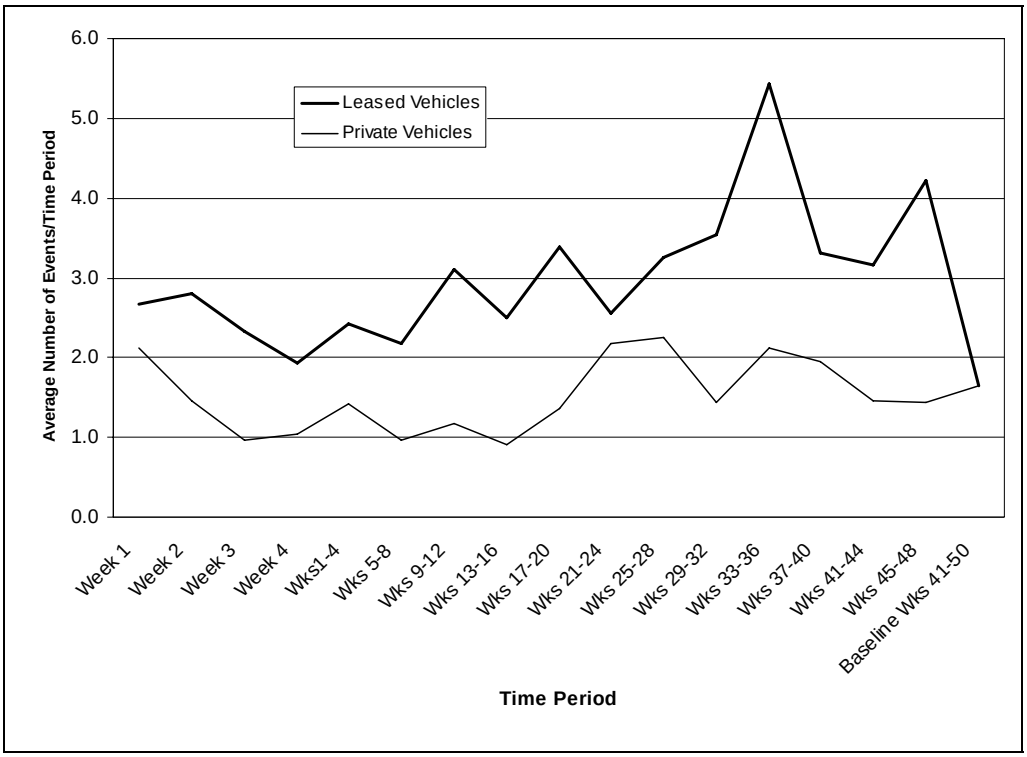


Figure 8.11. Mean number of events for younger drivers for leased and private vehicles for time periods of interest over weeks 1-50 of the study.

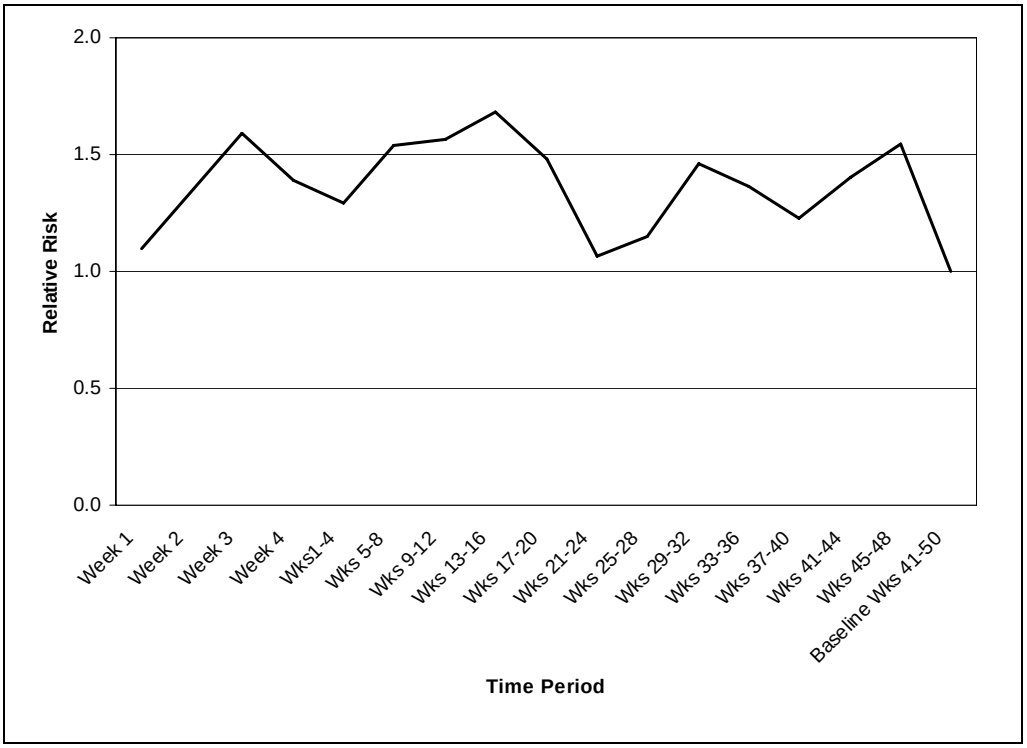


Figure 8.12. Relative risk for younger driver leased versus private vehicles for time periods of interest over weeks 1-50 of the study.

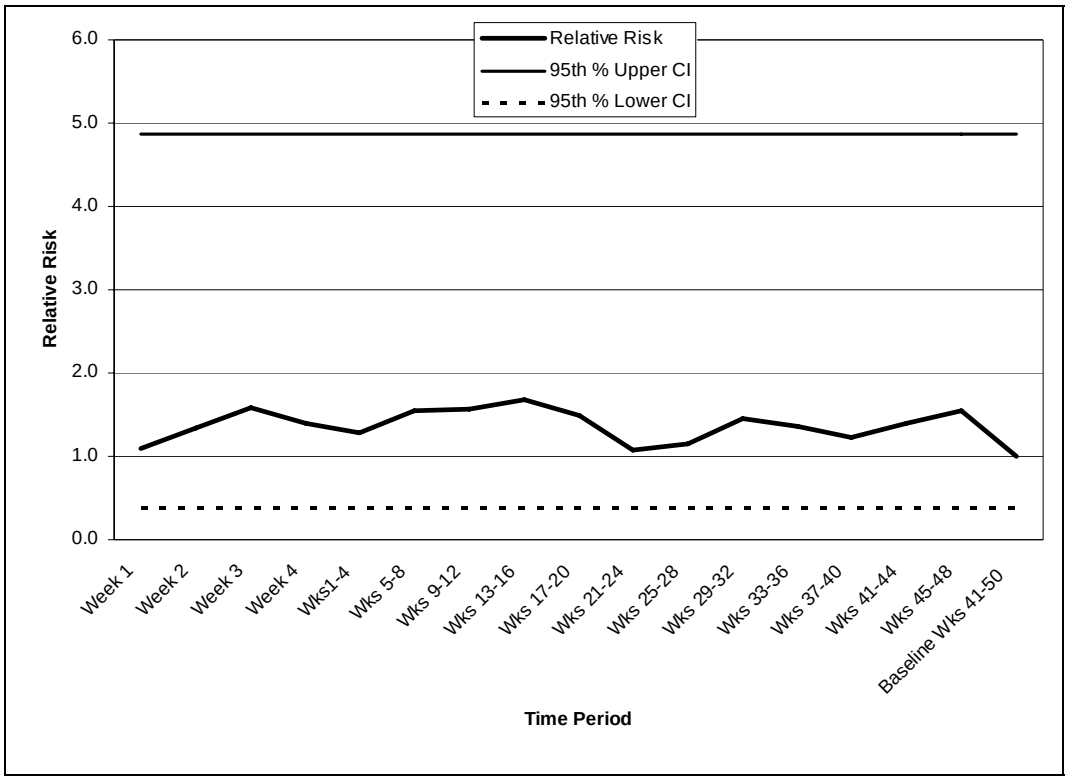


Figure 8.13. Relative risk and 95th percentage confidence intervals for younger driver leased versus private vehicles for time periods of interest over weeks 1-50 of the study.

Looking back at Figure 8.11 for the mean number of events, it is apparent that the leased versus private vehicle differences persist even when age is as carefully controlled as possible. This could indicate that there is a true difference in the relative risk between leased and private vehicles, and that it may have more to do with a lack of ownership of the vehicle than with age. Perhaps the participants were simply driving more recklessly with the leased vehicles because they would not have to pay for any scratches and dings that resulted from their carelessness. If this is true, you might expect that the difference would abate when only crashes and near-crashes are examined, because it seems unlikely that the drivers would be willing to put themselves in harm's way (as opposed to putting the vehicle in harm's way for an incident). In order to examine this hypothesis, the crashes and near-crashes were examined for the younger drivers of leased and private vehicles (again, controlling for age to the degree possible). Table 8.8 presents the means and RR for this younger driver near-crash analysis. Figure 8.14 presents the mean number of near-crashes for leased and private vehicles, Figure 8.15 presents the RR for near-crashes, and Figure 8.16 provides the confidence interval overlays.

Table 8.8. Mean number of near-crashes for younger drivers only per time period for leased and private vehicles; calculated RR for those time periods. Leased vehicle values do not include leased portion of switch drivers.

Time Period	Leased Average	Private Average	Relative Risk
Week 1	0.59	0.08	1.71
Week 2	0.48	0.21	1.17
Week 3	0.40	0.13	1.34
Week 4	0.15	0.09	1.22
Weeks 1-4	0.41	0.13	1.34
Weeks 5-8	0.24	0.11	1.26
Weeks 9-12	0.19	0.11	1.18
Weeks 13-16	0.24	0.07	1.56
Weeks 17-20	0.18	0.12	1.14
Weeks 21-24	0.25	0.17	1.10
Weeks 25-28	0.32	0.27	1.04
Weeks 29-32	0.34	0.07	1.68
Weeks 33-36	0.44	0.23	1.13
Weeks 37-40	0.24	0.12	1.25
Weeks 41-44	0.20	0.10	1.27
Weeks 45-48	0.31	0.07	1.63
Baseline Weeks 41-50	0.07	0.07	1.00
Weeks 1-50	0.28	0.13	1.25
Weeks 1-50 95% CI: Upper, Lower		52.07	0.030

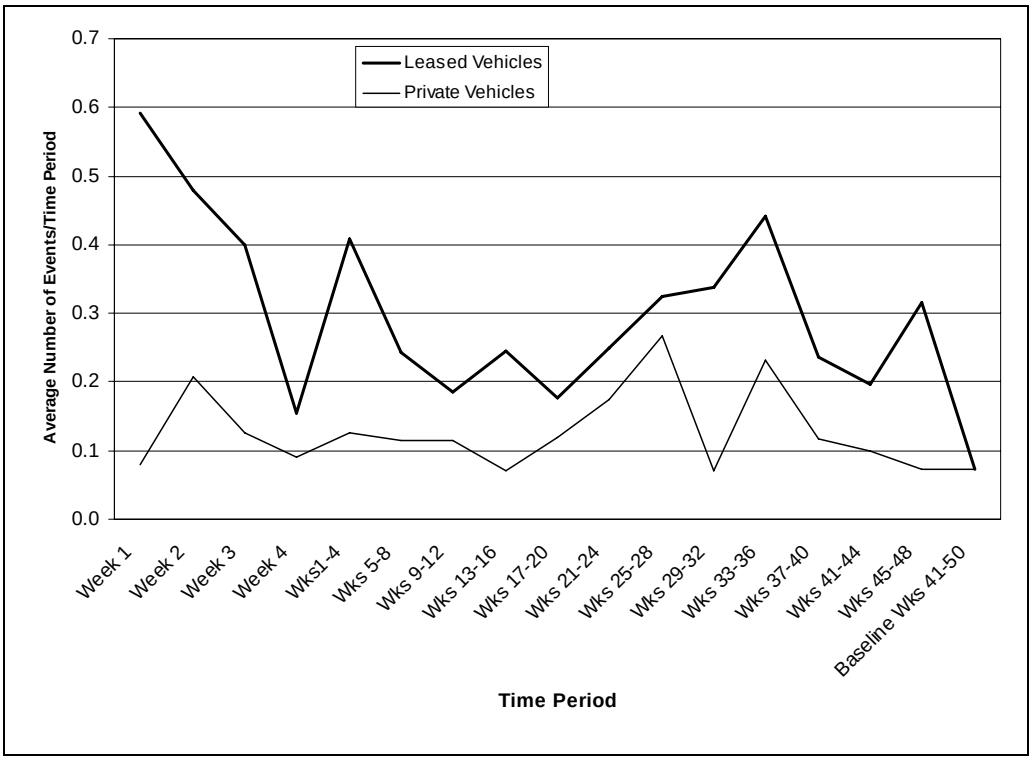


Figure 8.14. Mean number of near-crashes for younger drivers for leased and private vehicles for time periods of interest over weeks 1-50 of the study.

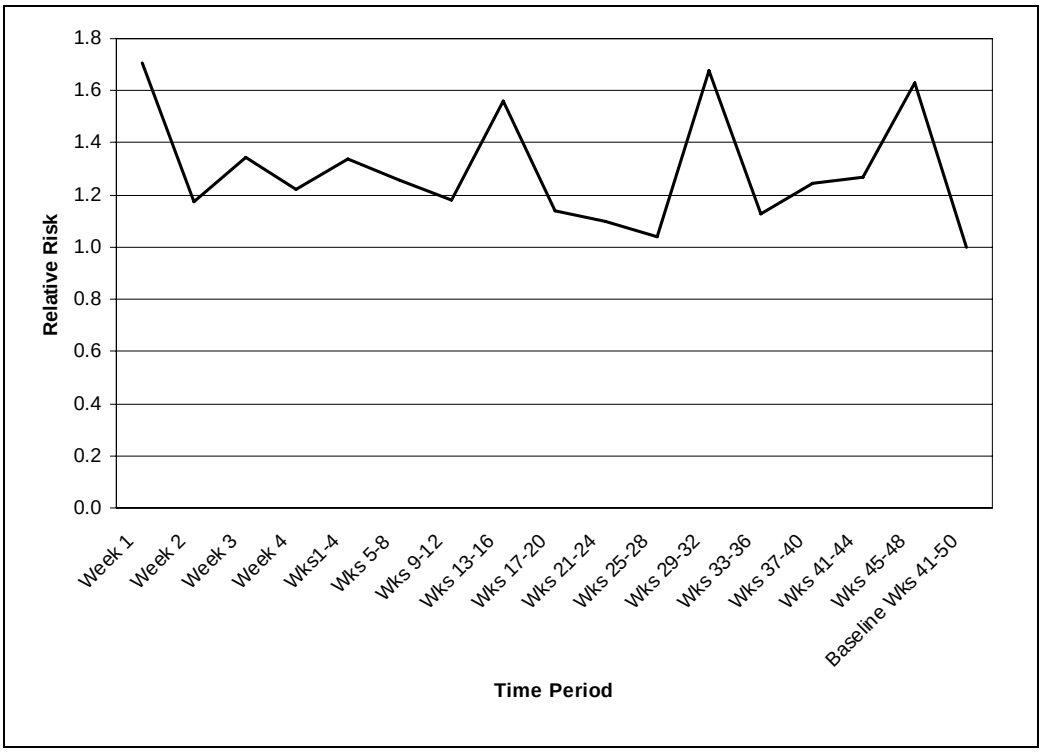


Figure 8.15. Relative risk for near-crashes for younger drivers leased versus private vehicles for time periods of interest over weeks 1-50 of the study.

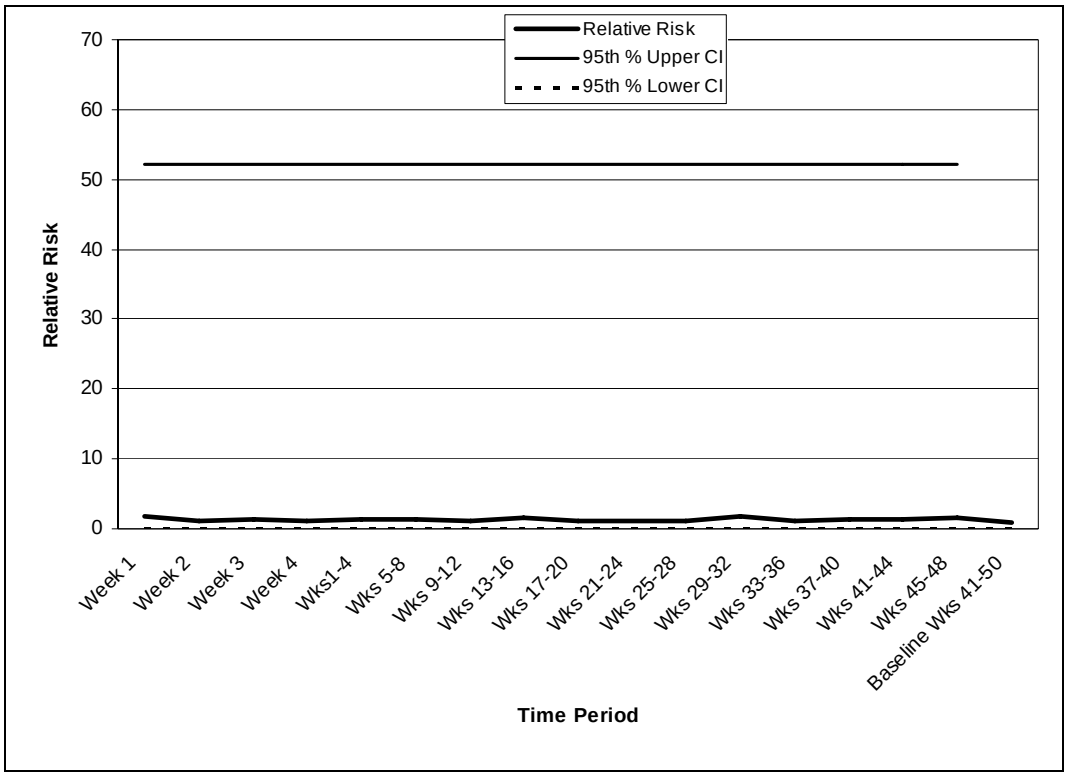


Figure 8.16. Relative risk and 95th percentage confidence intervals for near-crashes for younger drivers leased versus private vehicles for time periods of interest over weeks 1-50.

The hypothesis that leased vehicle drivers would not have a higher number of near-crashes was not borne out in this analysis. Note that the overall pattern for the mean number of near-crashes is still the same. The leased vehicles had a consistently higher number of near-crashes than the private vehicles. It is worth noting that the mean number of near-crashes was much higher for leased vehicles than for private vehicles in the first week of participation (0.59 and 0.08, respectively). This could be because the leased vehicle drivers were at greater risk because they were in an unfamiliar vehicle. This theme will be explored further in the hourly analyses.

The same sort of analysis was completed for the crashes as for the near-crashes, with the thought that even if the younger leased vehicle participants placed their vehicles in harm's way more often than private vehicle drivers, they might still have shown enough restraint to avoid getting in crashes. Table 8.9 shows the statistics for the crash analysis, while Figure 8.17 shows the mean number of crashes and Figure 8.18 shows the RR for crashes for leased versus private vehicles. The confidence intervals for crashes are now too large to include graphically.

Table 8.9. Mean number of crashes for younger drivers only per time period for leased and private vehicles; calculated RR for those time periods. Leased vehicle values do not include leased portion of switch drivers.

Time Period	Leased Average	Private Average	Relative Risk
Week 1	0.04	0.04	0.95
Week 2	0.04	0.00	NA
Week 3	0.04	0.00	NA
Week 4	0.04	0.05	0.90
Weeks1-4	0.04	0.02	1.55
Weeks 5-8	0.05	0.02	1.75
Weeks 9-12	0.03	0.01	2.02
Weeks 13-16	0.01	0.00	NA
Weeks 17-20	0.01	0.02	0.56
Weeks 21-24	0.03	0.01	2.01
Weeks 25-28	0.00	0.03	0.00
Weeks 29-32	0.04	0.00	NA
Weeks 33-36	0.08	0.06	1.16
Weeks 37-40	0.02	0.01	1.14
Weeks 41-44	0.00	0.01	0.00
Weeks 45-48	0.00	0.00	NA
Baseline Weeks 41-50	0.01	0.01	1.00
Weeks 1-50	0.03	0.02	1.35
Weeks 1-50 95% CI: Upper, Lower		94,968	0.00002

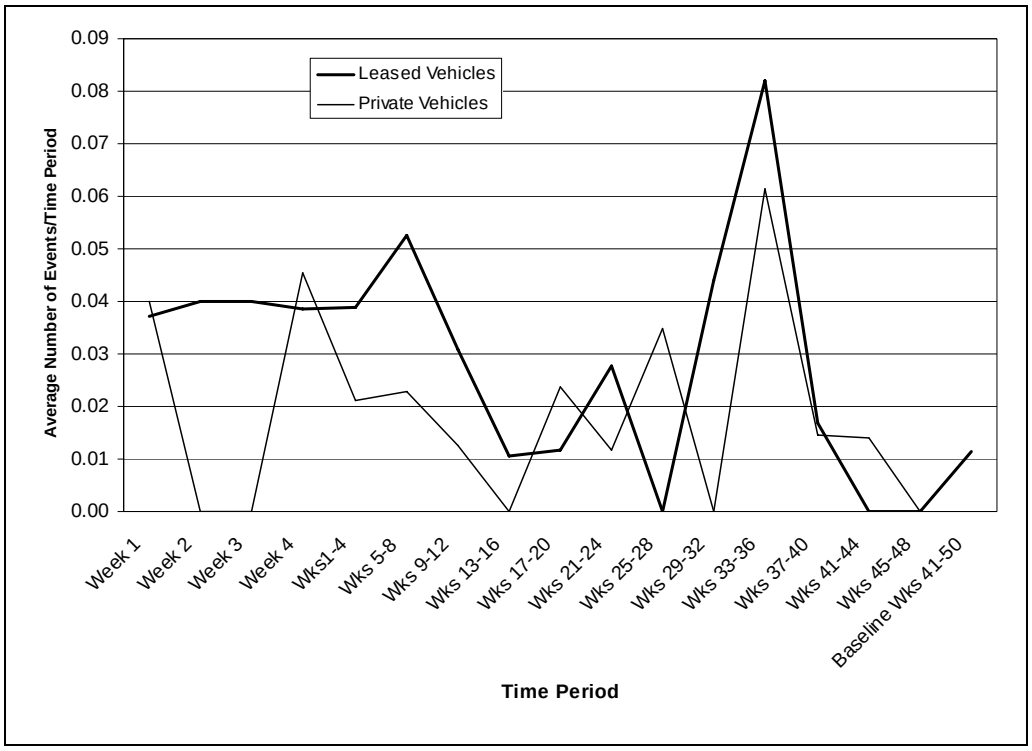


Figure 8.17. Mean number of crashes for younger drivers for leased and private vehicles for time periods of interest over weeks 1-50 of the study.

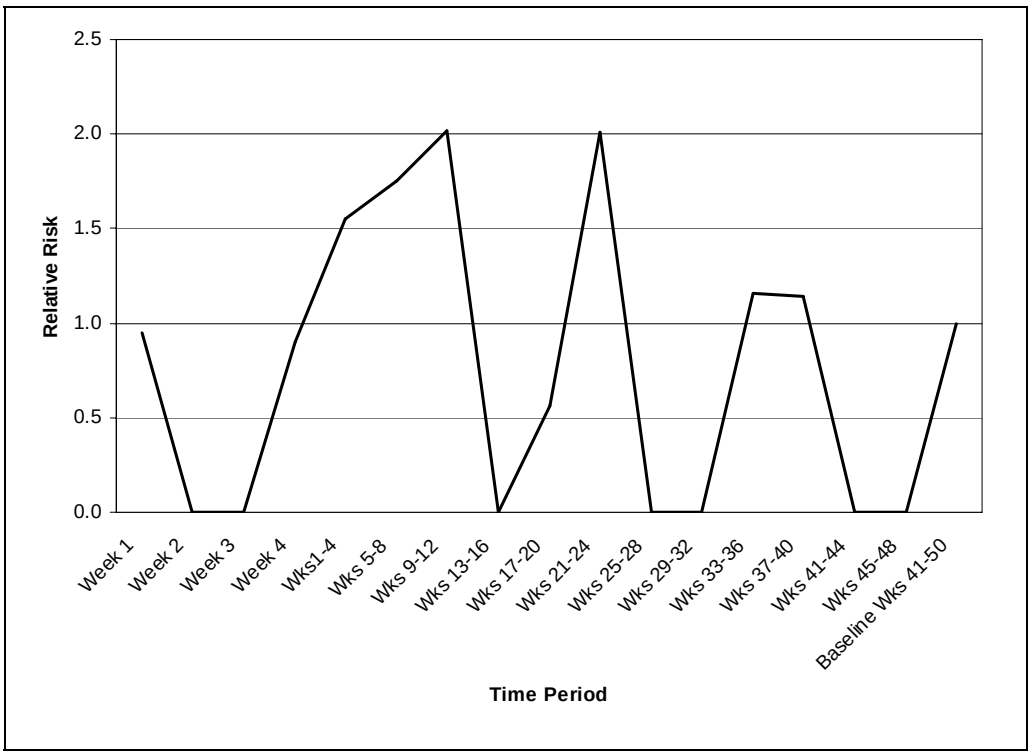


Figure 8.18. Relative risk for crashes for younger driver leased versus private vehicles for time periods of interest over weeks 1-50 of the study.

It can be seen in Figures 8.17 and 8.18 that the hypothesis regarding crash risk for leased vehicles does seem to hold true for crashes. Leased vehicle drivers seem to be willing to put their vehicles in harm's way, but not themselves. Based on these results, we might expect that if we were to transfer the leased vehicle drivers into their own private vehicles in which they would be responsible for repairs, insurance, etc., that their incident levels would drop to the same levels shown for private vehicle drivers. This question will be explored further in Questions 3 and 6 of Chapter 8, *Goal* 4. Altogether, there were 25 crashes for younger drivers in leased vehicles and 18 crashes for younger drivers in private vehicles in weeks 1 through 50. The next question will explore the hypothesis that drivers in a newly instrumented vehicle may experience a lower rate of events in their first hours of driving such a vehicle, which would then level off as they become used to the idea of instrumentation.

Question 2. Based on the number and type of valid events, is there a significant difference in the relative risk of driving over the first 50 hours for drivers in a vehicle with a newly installed instrumentation system?

Question 2 was designed to get at the idea of whether drivers experience an increase in valid events over the first few hours of driving a newly instrumented vehicle. The hypothesis was that drivers would not act naturally when they first began using an instrumented vehicle, and that they will begin to act more naturally as time goes on and they forget about the cameras and computers. It was hypothesized that the drivers would drive more carefully and experience fewer events when they were aware of the cameras, and that they would loosen their guard as time went on. If there were a point at which drivers adapted and began acting more naturally, this would be useful information for future instrumented vehicle studies of naturalistic driving. Previous experience at VTTI has indicated that drivers adapt amazingly quickly to the instrumented vehicle, but the question has never been empirically analyzed as will be attempted here. As seen in Question 1, we might expect to find differences in events between leased and private vehicles, even with a matched set of drivers, so these two groups of drivers will be explored, as they were in Question 1.

The time periods to be used in these analyses were hours 1, 2, 3, 4, and 5 (to check for differences in the first few hours), and hours 1-5, 6-10, 11-15, 16-20, 21-25, 26-30, 31-35, 36-40, 41-45, and 46-50. Hours 41-50 were averaged and used as the baseline time period. Note that there were 11 crashes in the first 50 hours of driving. In examining the weekly data, it was seen that there were 11 crashes in the first 5 weeks of driving. Therefore, 50 hours of driving was roughly equivalent to 5 weeks of driving, for an average of 10 hours of driving per week per driver. If the drivers did nearly all driving in 5 days, that would be an equivalent of 2 hours per day. If the 10 hours were spread over 7 days, the average amount of driving was 1.4 hours per day per driver.

The first question to be explored was whether there was a general downward trend in events across all drivers over the first 50 hours of driving. Figure 8.19 illustrates the mean number of events per hour for every hour, regardless of vehicle type or age. As can be seen, no clear trend was evident for all 974 events in the dataset. However, the lowest point was in the first hour, which gives some indication that there may be some differences between the first hour and the remaining hours.

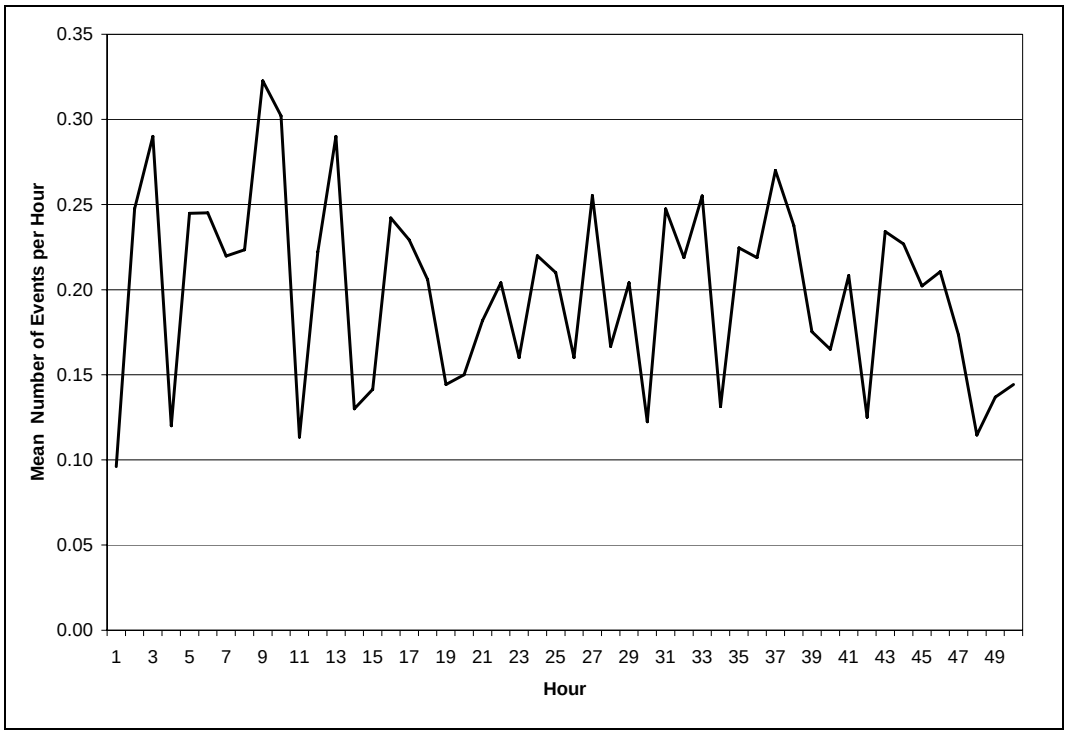


Figure 8.19. Overall trend of events over the first 50 hours of driving a newly instrumented vehicle.

The next analysis compares leased and private vehicles over the first 50 hours of driving, using the time periods of interest previously defined. Table 8.10 presents the means and RRs for these time periods. Figure 8.20 presents the mean number of events for private and leased vehicles, while Figure 8.21 shows the RR for these time periods. Due to the smaller number of events, the confidence intervals were larger, and were overlaid on any of the graphs in this section.

It can be seen in Figure 20 that the lowest mean numbers of events for private and leased vehicles were in hours 1 and 4. For every other time period, the leased vehicle mean exceeded the private vehicle mean, with no apparent trend for either vehicle type (alternating up and down, with an overall flat trend). This lends some credence to the hypothesis that drivers of both vehicle types were careful of the instrumentation system in the first hours, but very quickly acclimated and resumed a natural driving behavior.

Table 8.10. Mean number of events for per time period for leased and private vehicles; calculated RR for those time periods.

Time Period	Leased Average	Private Average	Relative Risk
Hour 1	0.07	0.09	0.87
Hour 2	0.36	0.14	1.46
Hour 3	0.38	0.20	1.26
Hour 4	0.08	0.12	0.77
Hour 5	0.30	0.18	1.23
Hours 1-5	0.23	0.15	1.24
Hours 6-10	0.33	0.18	1.26
Hours 11-15	0.22	0.13	1.28
Hours 16-20	0.32	0.11	1.61
Hours 21-25	0.22	0.13	1.27
Hours 26-30	0.25	0.14	1.29
Hours 31-35	0.24	0.14	1.28
Hours 36-40	0.35	0.14	1.48
Hours 41-45	0.22	0.16	1.16
Hours 46-50	0.22	0.13	1.31
Baseline Hours 41-50	0.16	0.16	1.00
Hours 1-50	0.26	0.14	1.31
Hours 1-50 95% CI: Upper, Lower		115	0.015

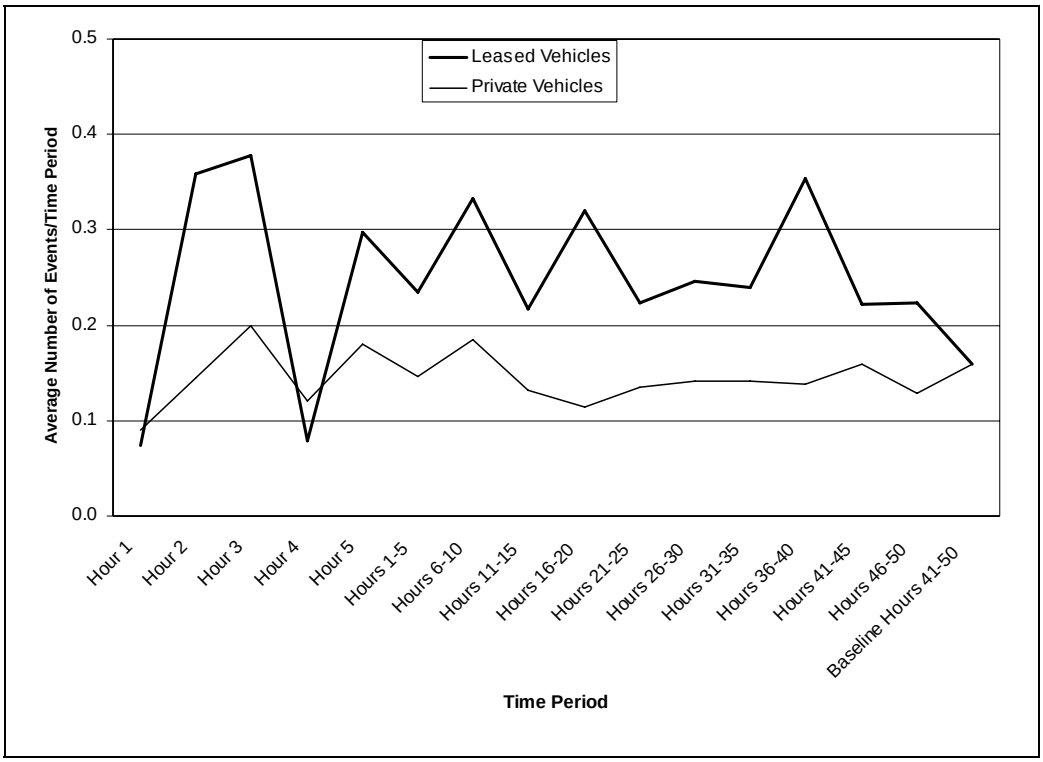


Figure 8.20. Mean number of events per vehicle for leased and private vehicles for time periods of interest over hours 1-50 of the study.

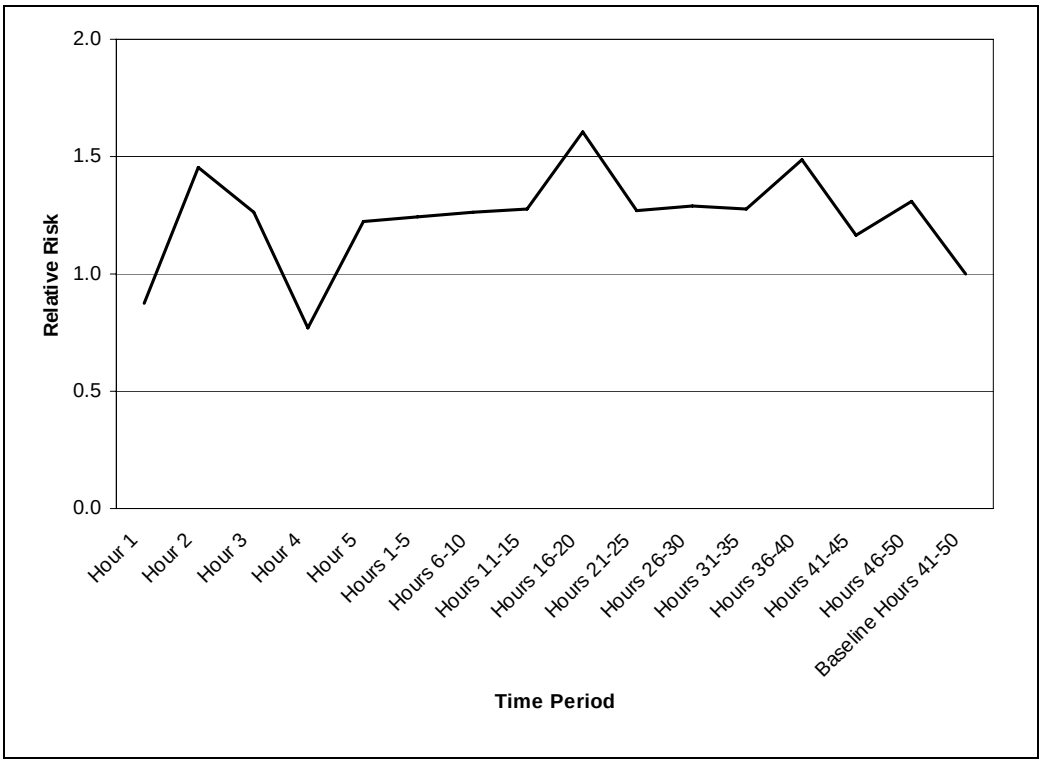


Figure 8.21. Leased versus private vehicle relative risk for time periods of interest over hours 1-50 of the study.

The next analysis used the matched set of younger drivers of leased and private vehicles in the same way that was done for Question 1. In this case, only overall events and near-crashes will be presented (there were only 5 leased vehicle younger driver crashes in the first 50 hours, and only 3 private vehicle crashes; the time periods for the crashes did not match up, so no RR analysis was possible for crashes). Table 8.11 presents the statistics, while Figure 8.22 shows the mean number of events and Figure 8.23 shows the RR for this analysis.

As before, even when controlling for age to the degree possible, leased vehicles experienced a greater mean number of events for nearly every time period studied. The only exceptions are hours 1 and 4, in which the leased and private vehicles experienced nearly identical mean numbers of events. Again, it appears that drivers of both vehicle types were being very careful during the first hour with a newly instrumented vehicle. Although not shown here, the age analysis had the same trend -- both younger and older drivers experienced their lowest mean rate of events in the first hour of driving, and both then leveled off to about the same levels seen in the weekly analysis.

Table 8.11. Mean number of events per time period for younger driver leased and private vehicles; calculated RR for those time periods.

Time Period	Leased Average	Private Average	Relative Risk
Hour 1	0.07	0.09	0.87
Hour 2	0.36	0.14	1.46
Hour 3	0.38	0.20	1.26
Hour 4	0.08	0.12	0.77
Hour 5	0.30	0.18	1.23
Hours 1-5	0.23	0.15	1.24
Hours 6-10	0.33	0.18	1.26
Hours 11-15	0.22	0.13	1.28
Hours 16-20	0.32	0.11	1.61
Hours 21-25	0.22	0.13	1.27
Hours 26-30	0.25	0.14	1.29
Hours 31-35	0.24	0.14	1.28
Hours 36-40	0.35	0.14	1.48
Hours 41-45	0.22	0.16	1.16
Hours 46-50	0.22	0.13	1.31
Baseline Hours 41-50	0.16	0.16	1.00
Hours 1-50	0.26	0.14	1.31
Hours 1-50 95% CI: Upper, Lower		92	0.020

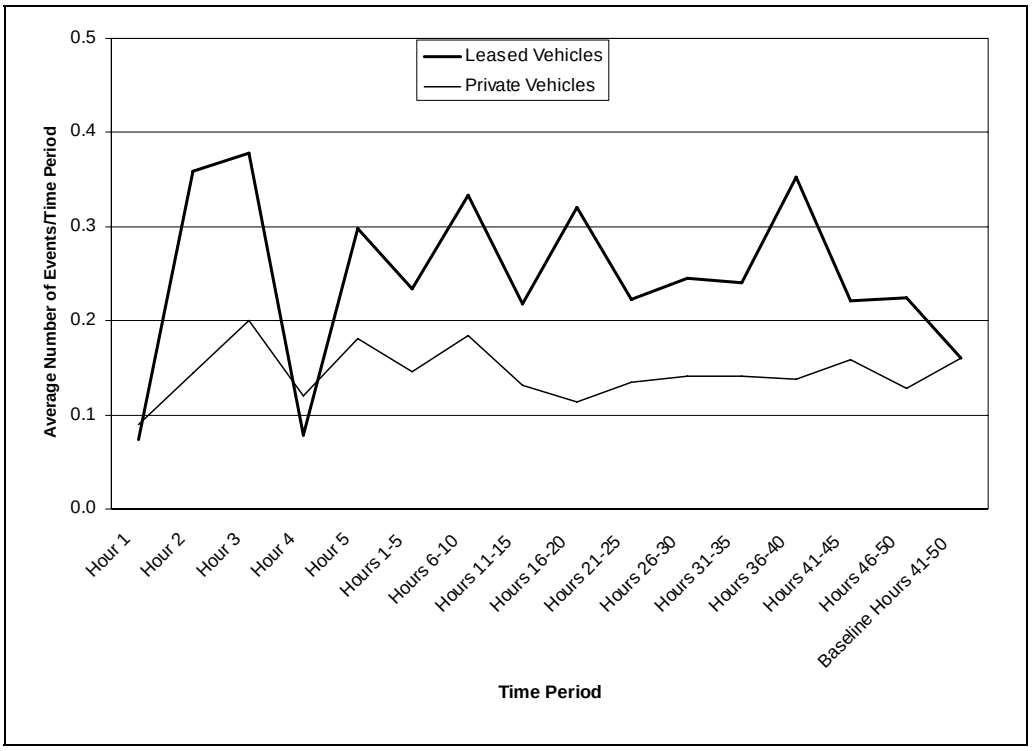


Figure 8.22. Mean number of events per vehicle for younger drivers for leased and private vehicles for time periods of interest over hours 1-50 of the study.

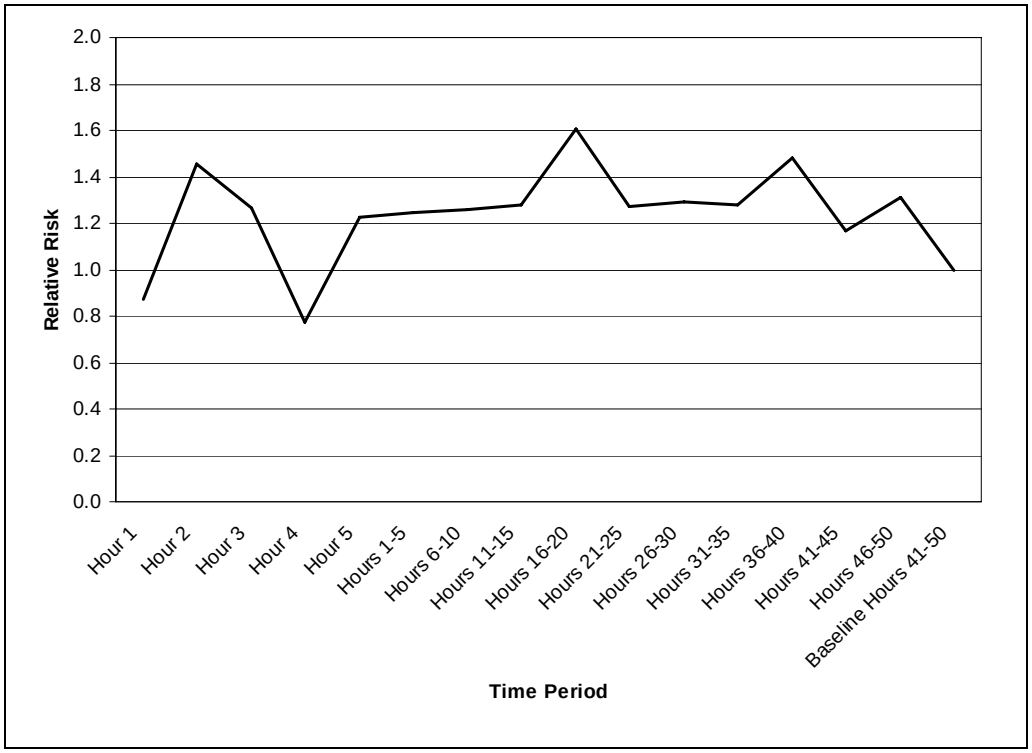


Figure 8.23. Younger driver leased versus private vehicle relative risk for time periods of interest over hours 1-50 of the study.

The final analysis was for near-crashes. There were 45 leased vehicle near-crashes and 22 private vehicle near-crashes for younger drivers in the first 50 hours of driving. The pattern was pretty much the familiar one, although there did seem to be an overall decline in near-crashes for leased vehicle drivers over the first 50 hours of driving, while the private vehicle driver near-crash levels remained nearly flat over the 50 hours. For nearly every time period, however, there were a greater mean number of near-crashes for the leased vehicle drivers as compared to the private vehicle drivers, even in this matched set of younger drivers. Table 8.12 presents the statistics, while Figure 8.24 shows the mean number of events and Figure 8.25 shows the RR for this analysis.

Table 8.12. Mean number of near-crashes per time period for younger driver leased and private vehicles; calculated RR for those time periods.

Time Period	Leased Average	Private Average	Relative Risk
Hour 1	0.02	0.01	1.14
Hour 2	0.05	0.01	1.24
Hour 3	0.05	0.00	NA
Hour 4	0.00	0.00	NA
Hour 5	0.08	0.00	NA
Hours 1-5	0.04	0.01	1.69
Hours 6-10	0.06	0.00	2.44
Hours 11-15	0.02	0.01	1.30
Hours 16-20	0.02	0.01	1.37
Hours 21-25	0.02	0.01	1.11
Hours 26-30	0.03	0.00	2.28
Hours 31-35	0.01	0.01	1.11
Hours 36-40	0.02	0.01	1.24
Hours 41-45	0.02	0.01	1.47
Hours 46-50	0.02	0.00	2.16
Baseline Hours 41-50	0.00	0.00	1.00
Hours 1-50	0.03	0.01	1.50
Hours 1-50 95% CI: Upper, Lower		92	0.020

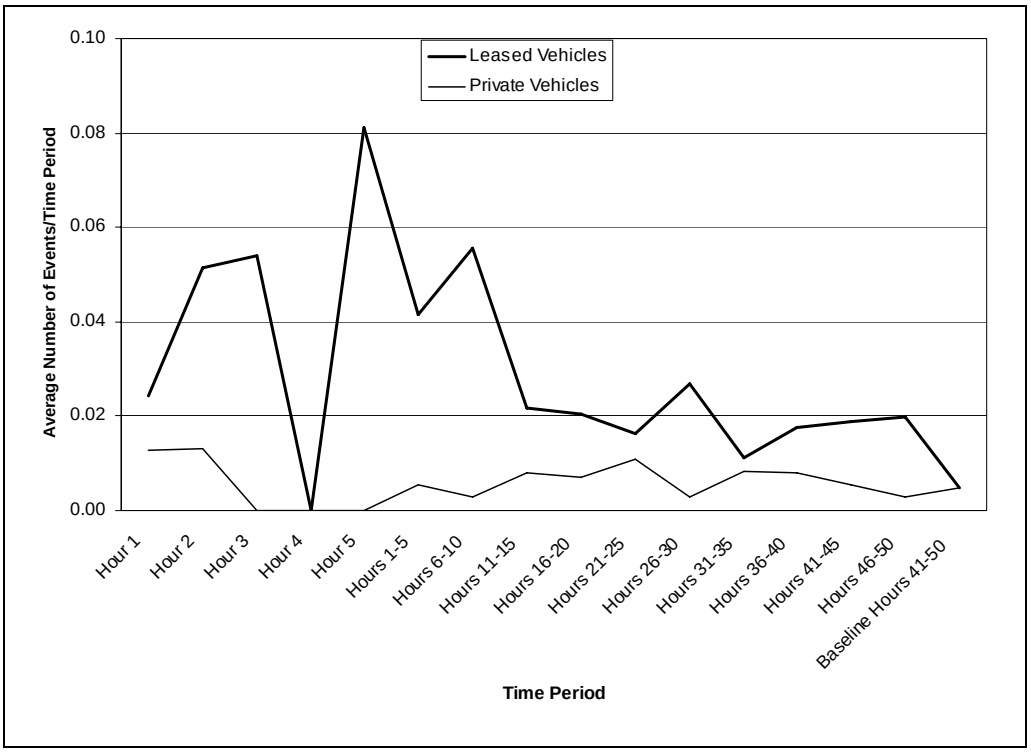


Figure 8.24. Mean number of near-crashes per vehicle for younger drivers for leased and private vehicles for time periods of interest over hours 1-50 of the study.

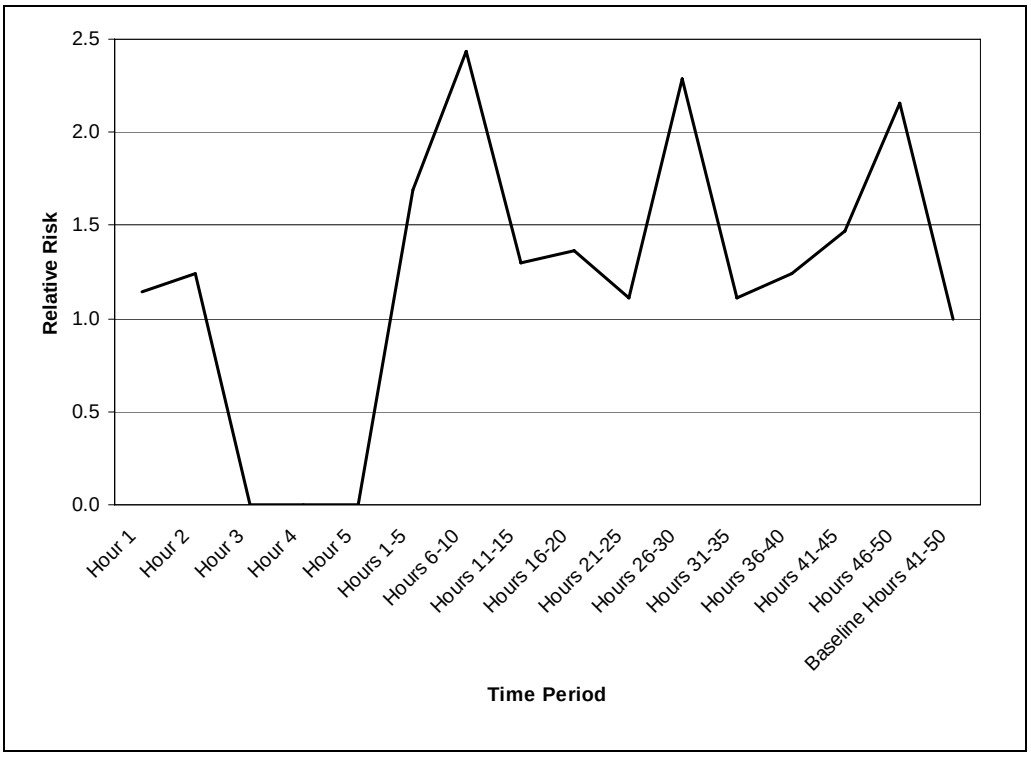


Figure 8.25. Younger driver leased versus private vehicle near-crash relative risk for time periods of interest over hours 1-50 of the study.

The analyses provided in this section provided some support for the thesis that drivers were more careful when first using an instrumented vehicle. The effect appears to wear off after the first hour. The dataset did not provide a breakdown by minutes, so it was not possible to tell whether this occurred within the first 5 minutes, the first half hour, or whether the adaptation did indeed happen after exactly one hour.

Question 3. Based on the number and type of valid events, is there a significant difference in the relative risk of driving for the same driver for weeks 1 through 4 of driving a privately owned vehicle and weeks 1 through 4 of driving a leased vehicle?

The purpose of this question was to investigate the driver adaptation process for the same driver in a leased vehicle versus a privately owned vehicle (both instrumented). Table 8.13 presents the data available for these analyses in terms of number of drivers and number of events. Only switch drivers for whom matched data were available for each week were used. For example, if there were no data for Driver 405 for Week 2 of leased vehicle driving, then the private vehicle driving for Driver 405 for Week 2 was also discarded. This resulted in a perfectly matched set of drivers for each week. The average number of events as well as the RR is also presented in Table 8.13. Weeks 2 through 4 were used as baseline for the RR calculations, and the leased average was set equal to the baseline average to provide for a control condition as was done in previous analyses. Figure 8.26 shows the mean number of events for leased and private vehicle driving for weeks 1-4, while Figure 8.27 presents the RR for these time periods. Figure 8.28 provides the 95th percentage confidence intervals for the RR. As shown in Figure 8.26, there were a greater mean number of events for leased vehicle driving on a week-by-week basis, even with a perfectly matched set of leased vehicle drivers.

Table 8.13. Statistics for a matched set of all switch drivers for weeks 1-4.

Week	Leased Drivers	Leased Events	Leased Average	Private Drivers	Private Events	Private Average	RR
Week 1	13	36	2.77	13	24	1.85	1.26
Week 2	11	48	4.36	11	32	2.91	1.21
Week 3	10	28	2.80	10	22	2.20	1.14
Week 4	9	37	4.11	9	13	1.44	1.78
Weeks 2-4 Baseline	30	113	3.03	30	91	3.03	1.00

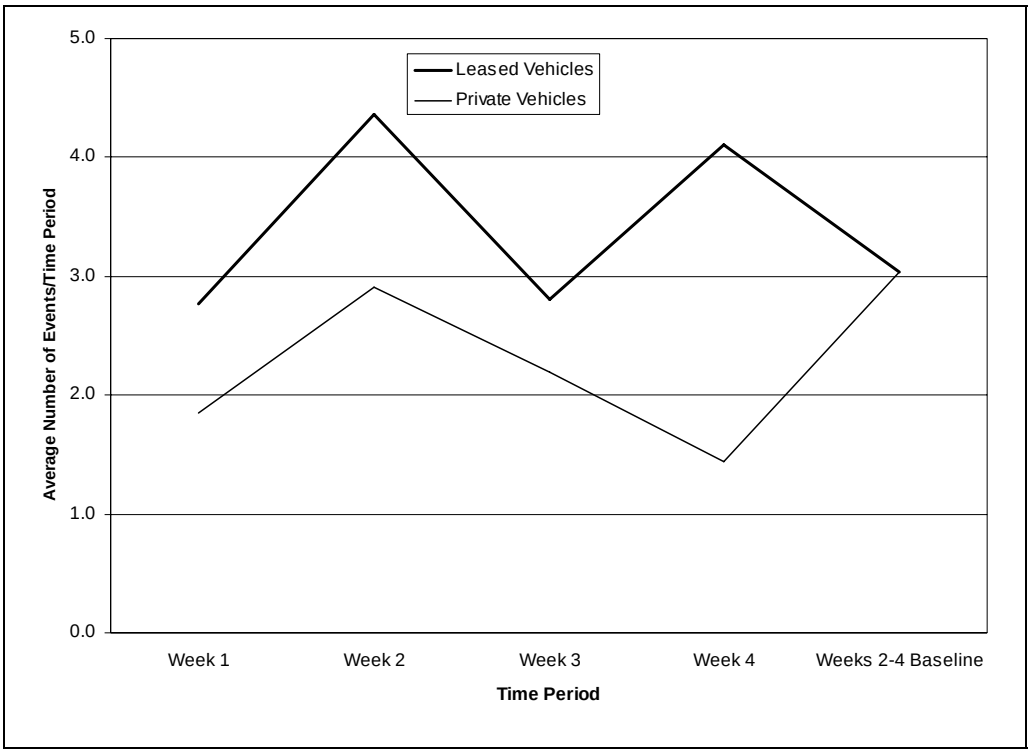


Figure 8.26. Matched set of switch drivers leased versus private vehicle mean number of events for weeks 1-4.

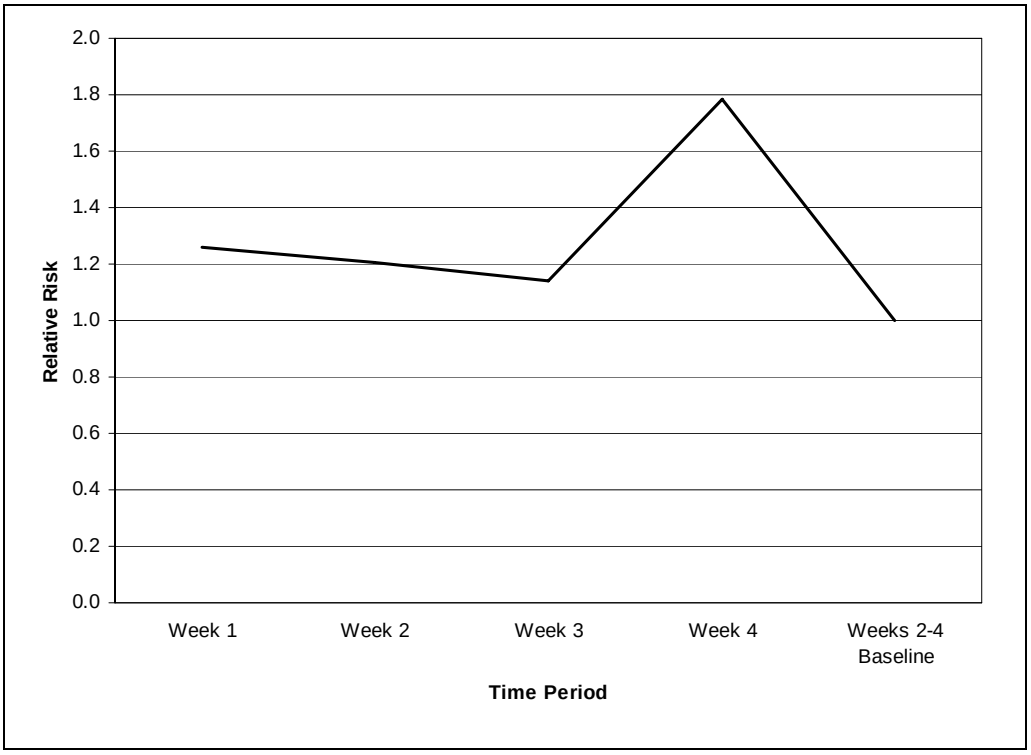


Figure 8.27. Matched set of switch drivers leased versus private vehicle relative risk for weeks 1-4.

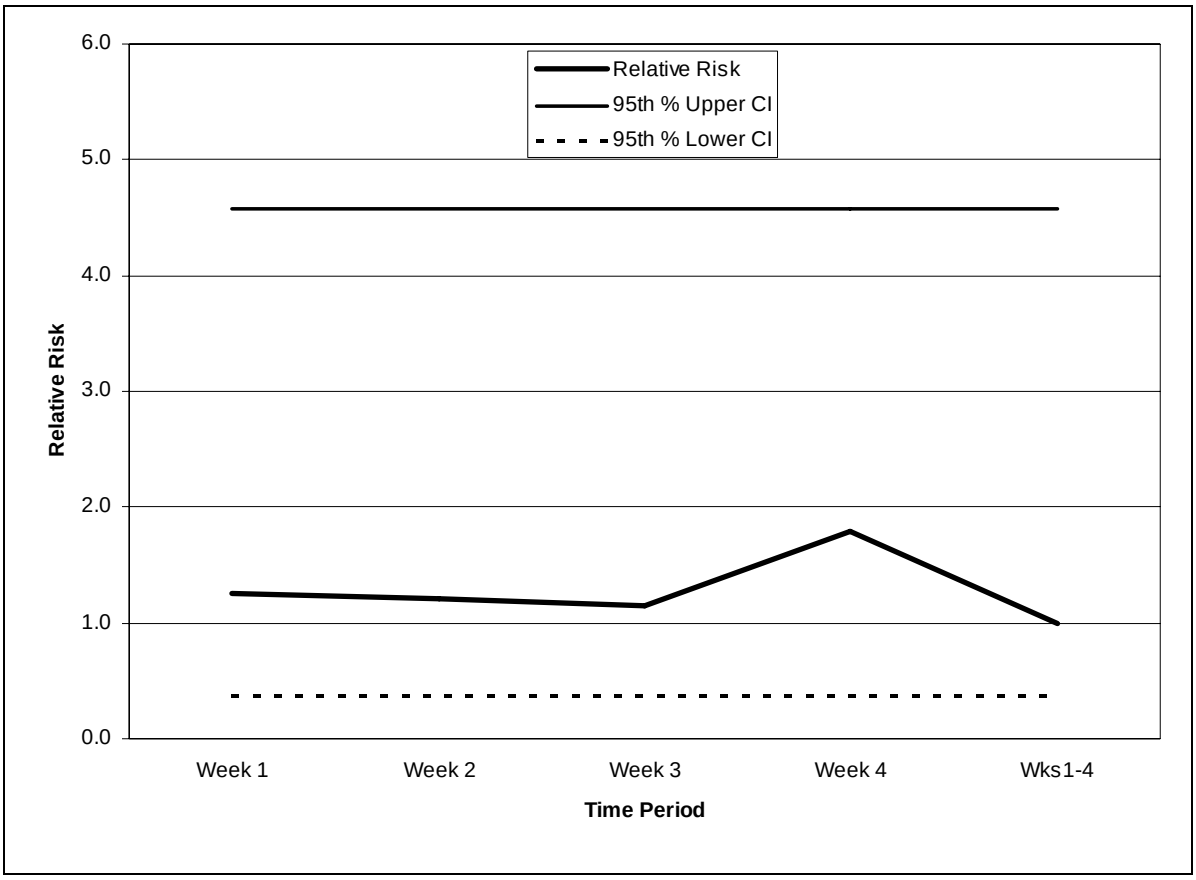


Figure 8.28. Matched set of switch drivers leased versus private vehicle relative risk for weeks 1-4 shown with 95th percentage upper and lower confidence intervals.

The relative risk for a matched set of younger switch drivers was examined next. Table 8.14 presents the statistics for the younger drivers. Figure 8.29 provides the mean values, Figure 8.30 shows the RR, and Figure 8.31 shows the RR with the confidence interval overlay. For the younger switch drivers, it appears that there was a slight downward trend for both the leased and private vehicles, but for every week, the leased vehicle driving had a higher mean rate of events than for private vehicle driving.

Table 8.14. Statistics for a matched set of all younger switch drivers for weeks 1-4.

Week	Leased Drivers	Leased Events	Leased Average	Private Drivers	Private Events	Private Average	RR
Week 1	4	12	3.00	4	9	2.25	1.10
Week 2	4	19	4.75	4	4	1.00	1.79
Week 3	4	14	3.50	4	6	1.50	1.36
Week 4	3	10	3.33	3	4	1.33	1.41
Weeks 2-4 Baseline	11	43	1.27	11	14	1.27	1.00

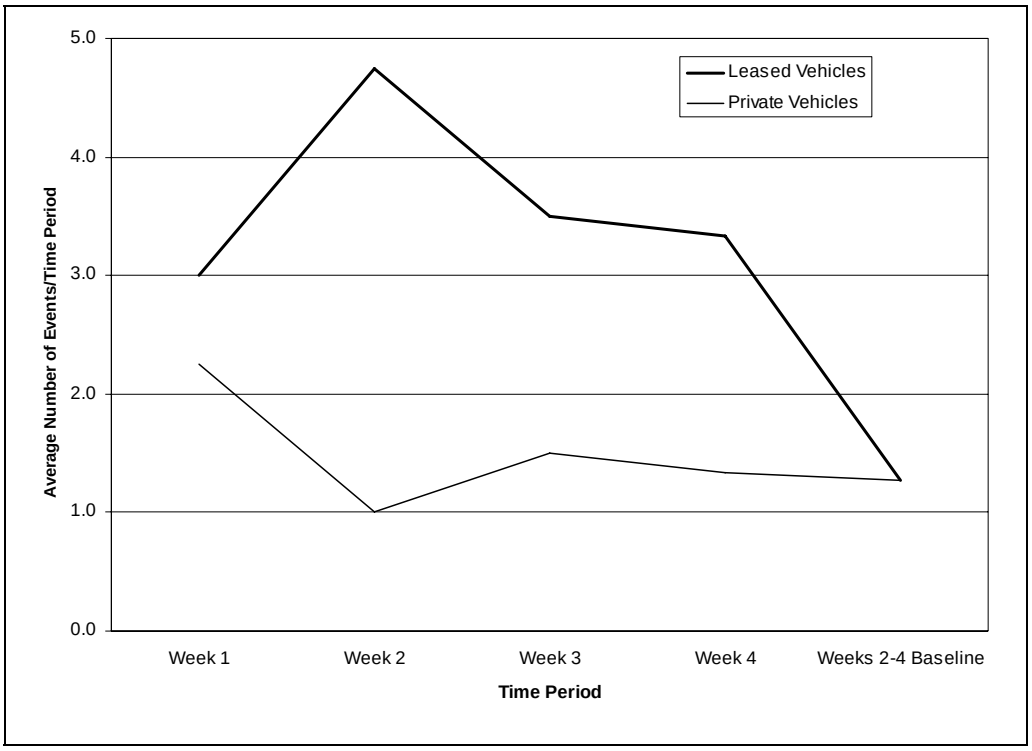


Figure 8.29. Matched set of younger switch drivers: leased versus private vehicle mean number of events for weeks 1-4.

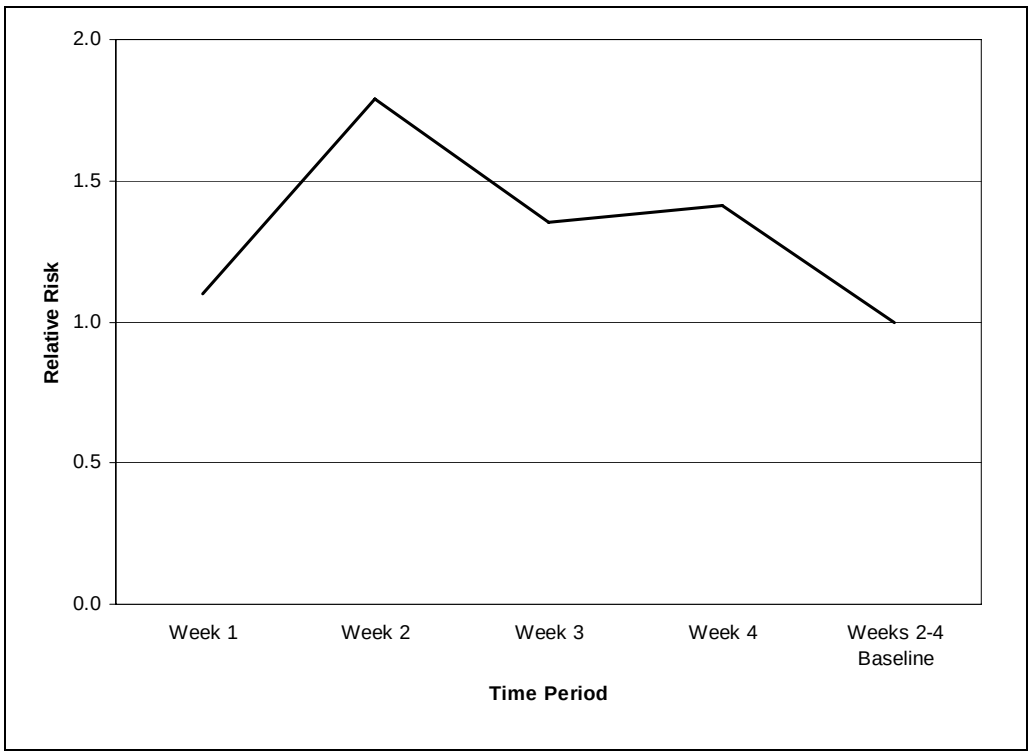


Figure 8.30. Matched set of younger switch drivers: leased versus private vehicle relative risk for weeks 1 -4.

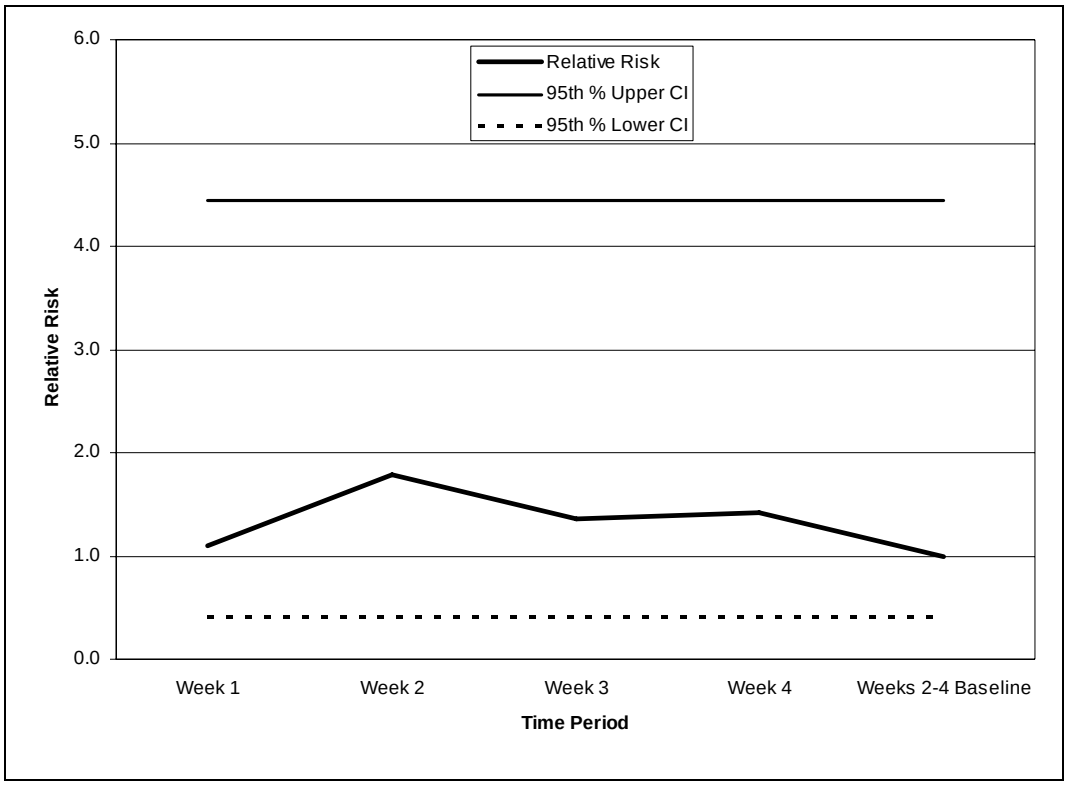


Figure 8.31. Matched set of younger switch drivers: leased versus private vehicle relative risk for weeks 1-4 shown with 95th percentage upper and lower confidence intervals.

The final analysis for Question 3 examines the matched set of older switch drivers. Table 8.15 provides the statistics, with the mean number of events for leased and private vehicles shown in Figure 8.32. Figure 8.33 illustrates the RR and Figure 8.34 provides the confidence intervals overlaid on the RR. This analysis was the first to show a clear break from the pattern seen up to this point − older switch drivers appeared to have virtually the same mean number of events for weeks 1 through 3, and then the usual difference between leased and private vehicles appeared.

Table 8.15. Statistics for a matched set of all older switch drivers for weeks 1-4.

Week	Leased Drivers	Leased Events	Leased Average	Private Drivers	Private Events	Private Average	RR
Week 1	9	24	2.67	9	15	1.67	1.31
Week 2	7	29	4.14	7	28	4.00	1.01
Week 3	6	14	2.33	6	16	2.67	0.93
Week 4	6	27	4.50	6	9	1.50	1.77
Weeks 2-4 Baseline	19	70	2.79	19	53	2.79	1.00

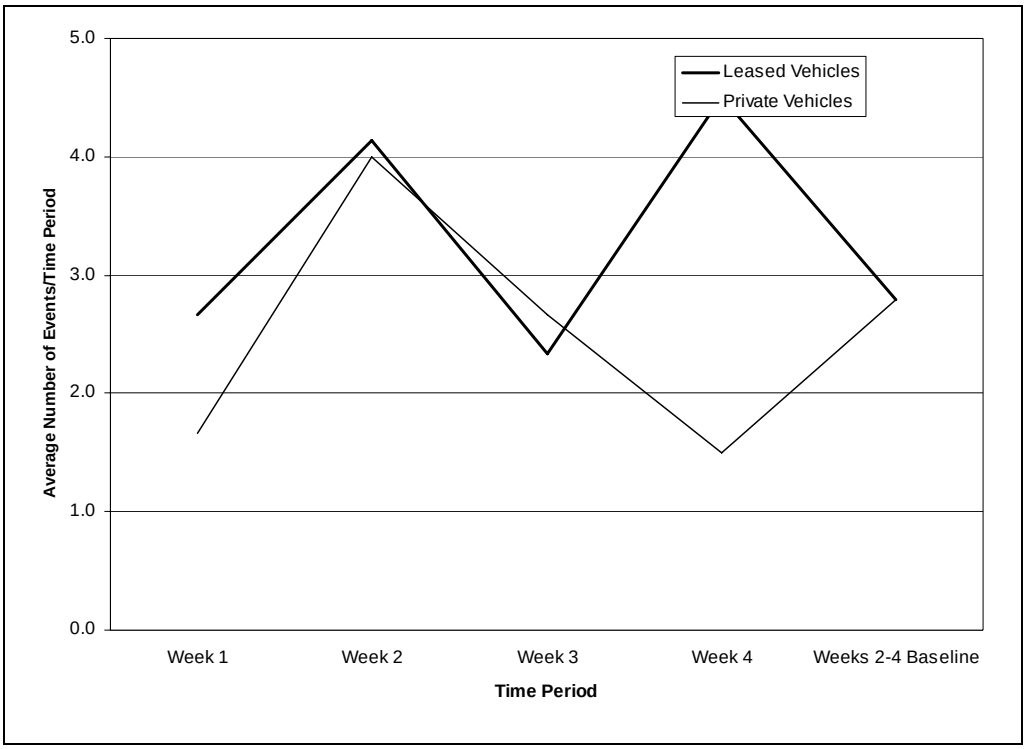


Figure 8.32. Matched set of older switch drivers: leased versus private vehicle mean number of events for weeks 1-4.

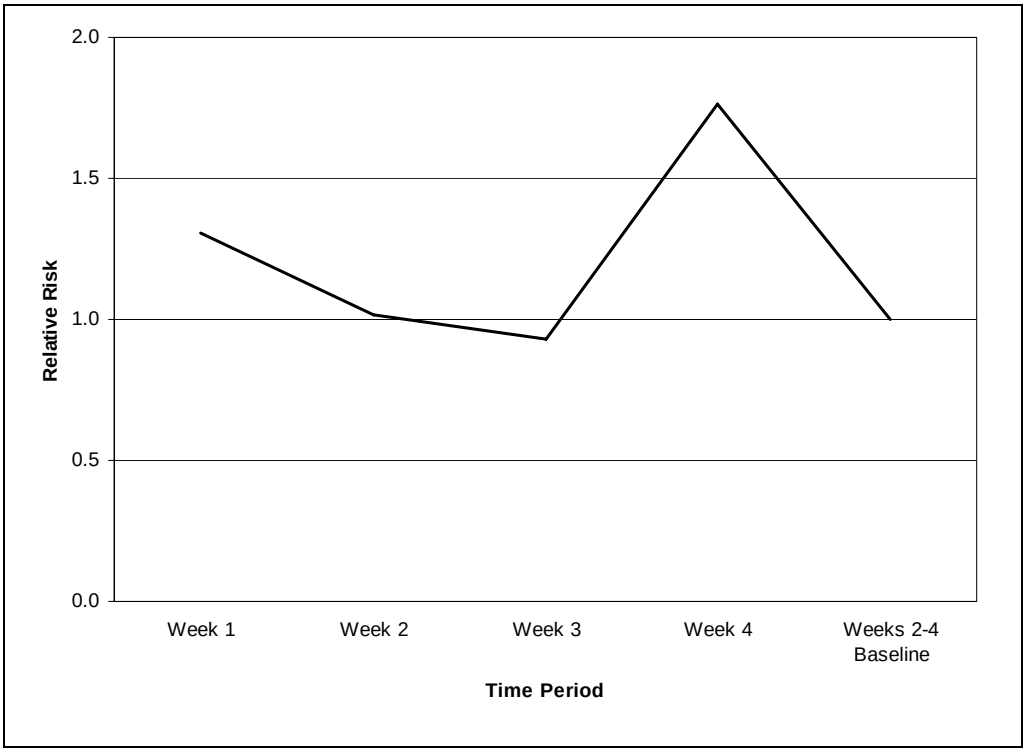


Figure 8.33. Matched set of older switch drivers: leased versus private vehicle relative risk for weeks 1-4.

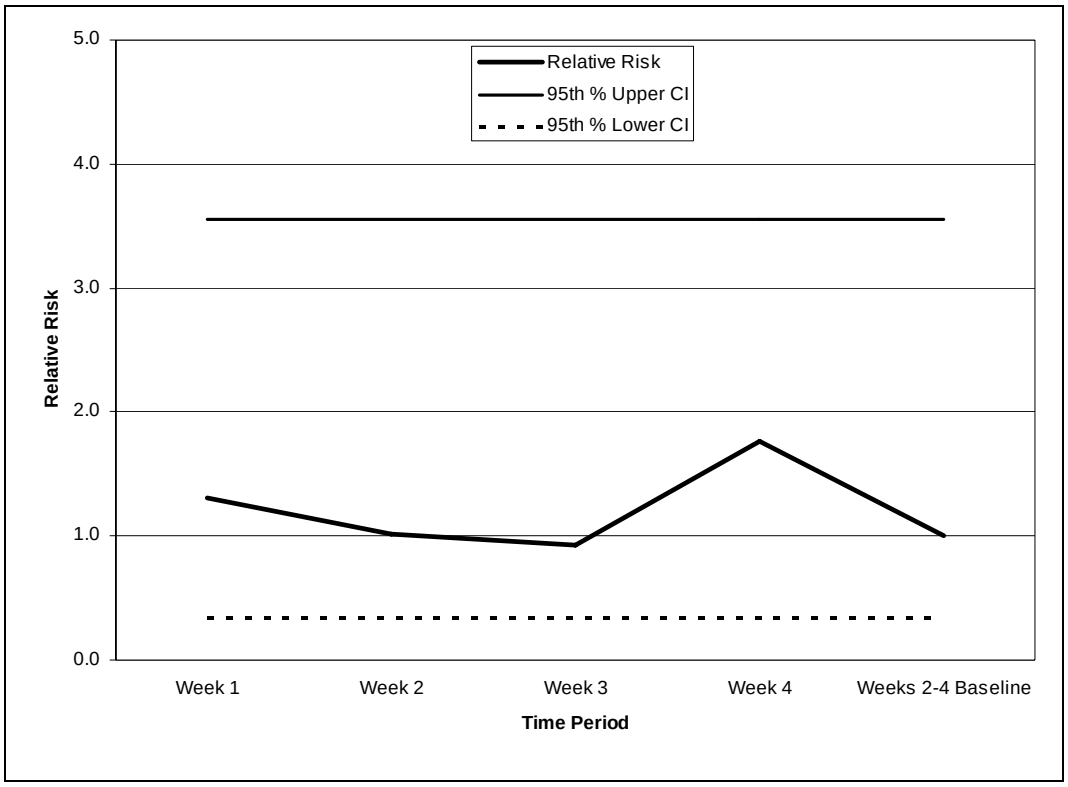


Figure 8.34. Matched set of older switch drivers: leased versus private vehicle relative risk for weeks 1-4 shown with 95th percentage upper and lower confidence intervals.

The analyses using perfectly matched sets of switch drivers had similar results to the previous analyses. Even when the same driver was switched from a private vehicle to a leased vehicle, there were still more events in the leased vehicle than in the private vehicle. There was some indication that older switch drivers might not fit this profile for weeks 1 through 3, but by week 4, the difference in the mean number of events between leased and private vehicles was back at the usual level. If the increased number of events in leased vehicle driving for the same driver was due to vehicle unfamiliarity, this effect was not extinguished over the first four weeks.

There were only 25 near-crashes among the switch drivers in the first 4 weeks of driving, so no near-crash analyses were attempted for this question. The 25 near-crashes would have been split amongst the cells representing weeks 1 through 4 and leased and private driving, leaving very few data points in any given cell. Likewise, there were only 2 crashes in the dataset for this question.

As distinguished from the raw numbers used in Questions 1, 2, and 3, similar analyses will be performed to examine driver adaptation from the perspective of rate of valid events per mile driven for Questions 4, 5, and 6.

Question 4. Based on the rate of valid events per mile driven, is there a significant difference in the relative risk of driving over the course of a year for drivers in a familiar vehicle with instrumentation installed (leased and private vehicles)?

The purpose of this question was to investigate the driver adaptation process for leased vehicles and privately owned vehicles with newly installed instrumentation over the course of the study, while accounting for exposure by analyzing the number of events per mile driven. As before for weekly analysis, an average of weeks 41-50 was used as the control time period. Events per mile were calculated by dividing the number of valid events per week by the number of miles driven for that week. Table 8.16 presents the mean number of events per mile for private and leased vehicles for the time periods of interest as described previously. The mean number of events for private and leased vehicles for these time periods is presented in Figure 8.35, while the RR for these time periods is shown in Figure 8.36. As can be seen, the previous finding of a greater mean number of events for leased vehicles was continued even when corrected for exposure, although there was a slight upward trend for both leased and private vehicles. Because the magnitude of the difference between leased and private vehicles stayed nearly the same throughout the 50 weeks, the RR stayed close to 1 and fairly flat over the time periods of interest.

Table 8.16. Mean number of events per vehicle per mile for leased and private vehicles; calculated relative risk for that time period.

Time Period	Leased Average	Private Average	Relative Risk
Week 1	0.0033	0.0028	1.11
Week 2	0.0044	0.0034	1.15
Week 3	0.0042	0.0026	1.33
Week 4	0.0048	0.0027	1.35
Weeks1-4	0.0041	0.0029	1.22
Weeks 5-8	0.0036	0.0027	1.18
Weeks 9-12	0.0042	0.0029	1.23
Weeks 13-16	0.0042	0.0026	1.30
Weeks 17-20	0.0038	0.0027	1.22
Weeks 21-24	0.0040	0.0033	1.12
Weeks 25-28	0.0049	0.0036	1.17
Weeks 29-32	0.0060	0.0037	1.26
Weeks 33-36	0.0052	0.0038	1.17
Weeks 37-40	0.0047	0.0046	1.02
Weeks 41-44	0.0041	0.0041	1.00
Weeks 45-48	0.0071	0.0041	1.28
Baseline Weeks 141-50	0.0055	0.0043	1.00
Weeks 1-50	0.0045	0.0034	1.15

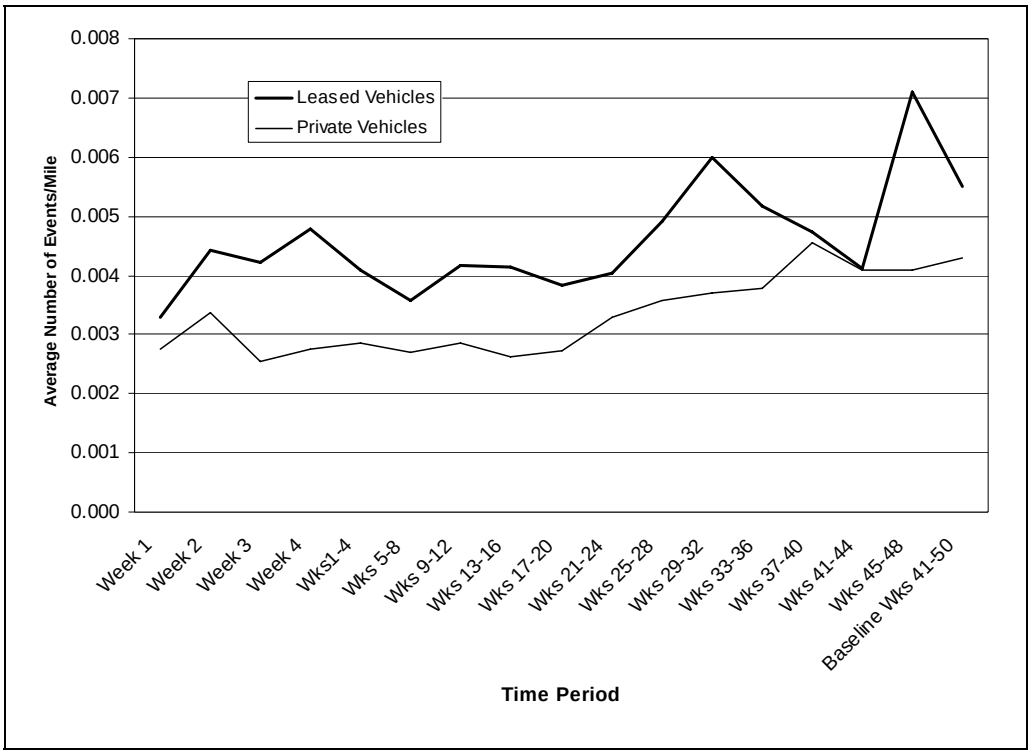


Figure 8.35. Mean number of events per mile for leased and private vehicles for time periods of interest over weeks 1-50 of the study.

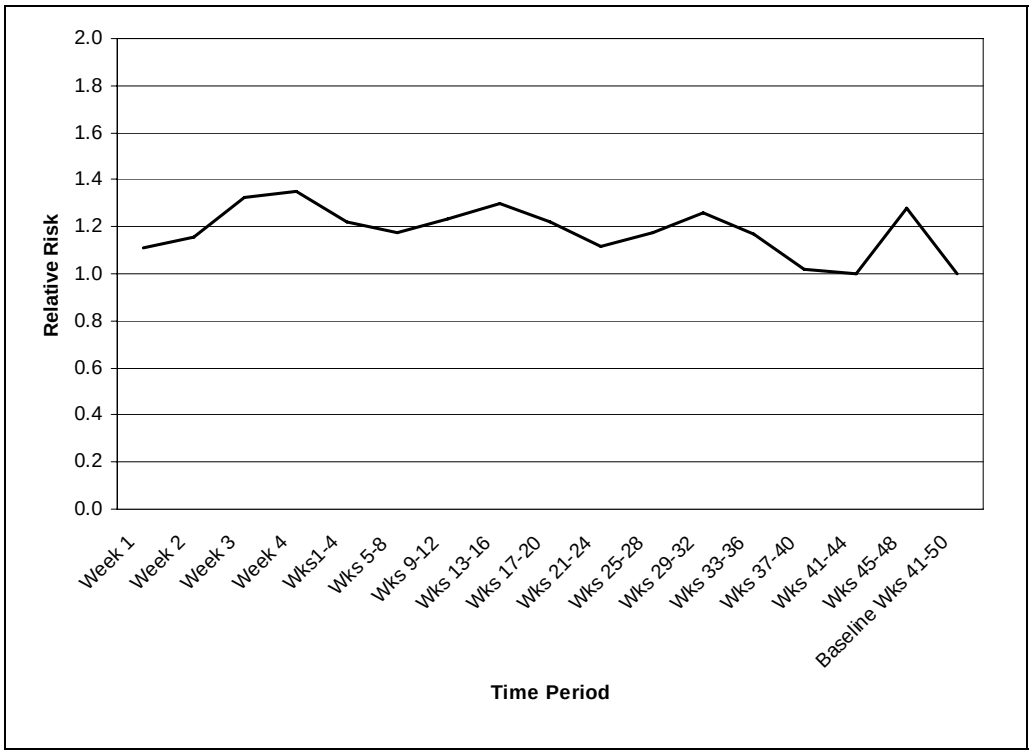


Figure 8.36. Events per mile relative risk for leased versus private vehicles for time periods of interest over weeks 1-50 of the study.

Given the slight upward trend of events per mile for leased and private vehicles, a matched set of younger drivers was used to determine whether this trend continued with a more controlled group of participants. Given that the numbers and ages were a reasonable matched set as shown in Question 1, a comparison of leased versus private vehicles was performed including only the younger drivers. Table 8.17 shows the means and RR for each time period of interest. Figure 8.37 shows the mean number of events per mile for each time period, while Figure 8.38 shows the RR for these time periods. As can be seen in Figure 8.37, the upward trend was even more pronounced for the matched set of younger drivers. In examining the miles and event raw data, it was apparent that fewer miles were driven (or recorded) as time went on, and that there was a less steep downward decline in the number of events over time. These two factors combined to create the increase in events per mile over time. Whether the decline in miles driven over time is accurate was difficult to determine, so this result may be an artifact of the data reduction process.

Table 8.17. Mean number of events per mile for younger drivers only for leased and private vehicles; calculated RR for those time periods. Leased vehicle values do not include leased portion of switch drivers.

Time Period	Leased Average	Private Average	Relative Risk
Week 1	0.0028	0.0030	0.95
Week 2	0.0034	0.0032	1.05
Week 3	0.0035	0.0021	1.39
Week 4	0.0040	0.0027	1.28
Weeks1-4	0.0034	0.0028	1.15
Weeks 5-8	0.0032	0.0024	1.21
Weeks 9-12	0.0043	0.0029	1.26
Weeks 13-16	0.0043	0.0025	1.39
Weeks 17-20	0.0040	0.0028	1.24
Weeks 21-24	0.0041	0.0034	1.12
Weeks 25-28	0.0049	0.0036	1.19
Weeks 29-32	0.0060	0.0040	1.25
Weeks 33-36	0.0053	0.0047	1.06
Weeks 37-40	0.0048	0.0048	1.01
Weeks 41-44	0.0046	0.0039	1.09
Weeks 45-48	0.0071	0.0048	1.20
Baseline Weeks 141-50	0.0054	0.0054	1.00
Weeks 1-50	0.0047	0.0038	1.13

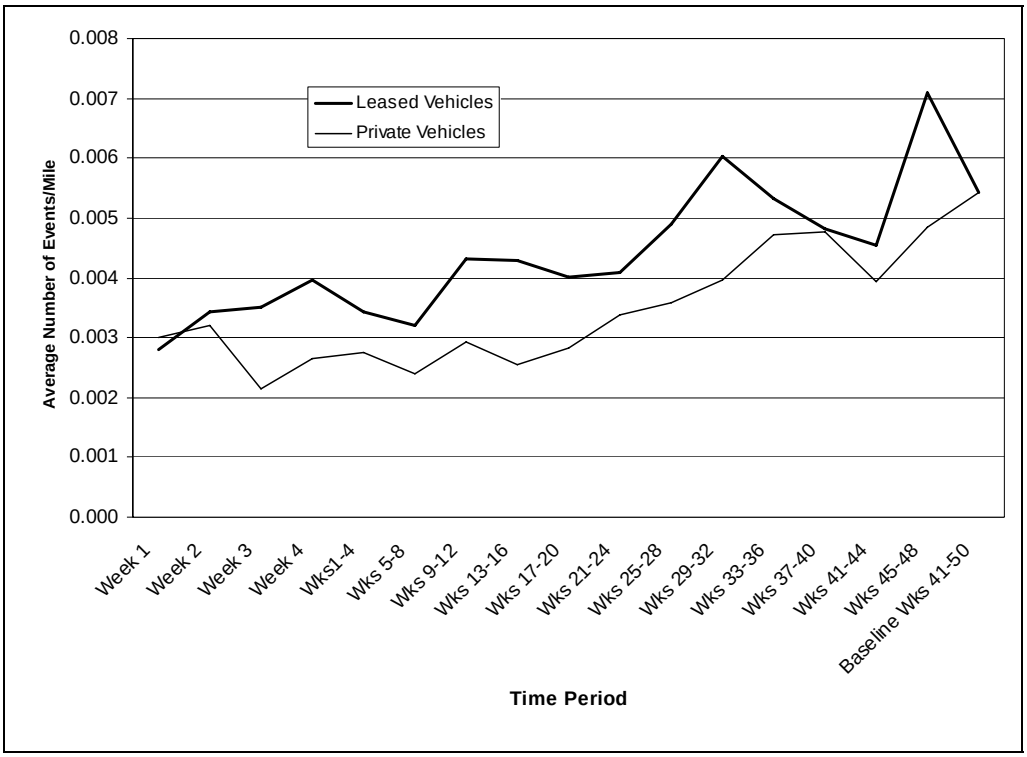


Figure 8.37. Mean number of events per mile for younger drivers for leased and private vehicles for time periods of interest over weeks 1-50 of the study.

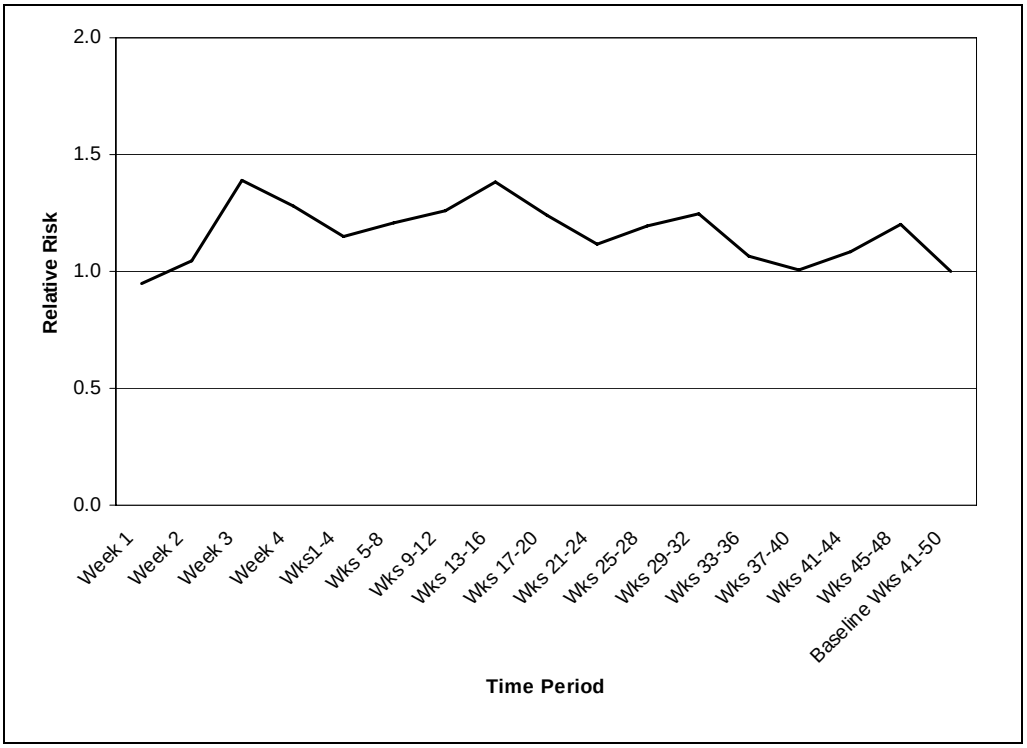


Figure 8.38. Relative risk for younger driver leased versus private vehicles for time periods of interest over weeks 1-50 of the study.

Question 5. Based on the number of valid events per mile, is there a significant difference in the relative risk of driving over the first 50 hours for drivers in a vehicle with a newly installed instrumentation system?

Question 5 was designed to get at the idea of whether drivers experienced an increase in valid events over the first few hours of driving a newly instrumented vehicle when exposure was controlled by examining the data on a per mile basis. The hypothesis was that drivers would not act naturally when they first began using an instrumented vehicle, and that they would begin to act more naturally as time went on and they would forget about the cameras and computers. It is hypothesized that the drivers would drive more carefully and experience fewer events when they were aware of the cameras, and that they would loosen their guard as time goes on. As with previous questions, we might expect to find differences in events between leased and private vehicles, even with a matched set of drivers, so these two groups of drivers will be explored.

The next analysis compared leased and private vehicles over the first 50 hours of driving, using the time periods of interest previously defined and calculating events per mile in the same way. Table 8.18 presents the means and RRs for these time periods. Figure 8.39 presents the mean number of events for private and leased vehicles, while Figure 8.40 shows the RR for these time periods.

It can be seen in Figure 8.39 that the data followed almost the exact same pattern as for Question 2. The first hour provided the lowest mean rate of events, and leased and personal vehicles were similar for the hours 1 and 4. Other than hours 1 and 4, the leased vehicle event per mile rate was higher than the rate for private vehicles for all other time periods. The magnitude of the difference was about the same as seen in previous analyses. This lends some credence to the hypothesis that drivers of both vehicle types were careful of the instrumentation system in the first hours, but very quickly acclimated and resumed a natural driving behavior.

Table 8.18. Mean number of events for per mile for leased and private vehicles; calculated RR for those time periods.

Time Period	Leased Average	Private Average	Relative Risk
Hour 1	0.0022	0.0030	0.80
Hour 2	0.0131	0.0064	1.36
Hour 3	0.0124	0.0091	1.12
Hour 4	0.0031	0.0049	0.75
Hour 5	0.0115	0.0077	1.18
Hours 1-5	0.0082	0.0060	1.17
Hours 6-10	0.0129	0.0081	1.20
Hours 11-15	0.0077	0.0058	1.16
Hours 16-20	0.0112	0.0050	1.46
Hours 21-25	0.0095	0.0060	1.25
Hours 26-30	0.0102	0.0062	1.25
Hours 31-35	0.0100	0.0060	1.27
Hours 36-40	0.0137	0.0059	1.44
Hours 41-45	0.0096	0.0067	1.18
Hours 46-50	0.0091	0.0053	1.31
Baseline Hours 41-50	0.0067	0.0067	1.00
Hours 1-50	0.0101	0.0061	1.26

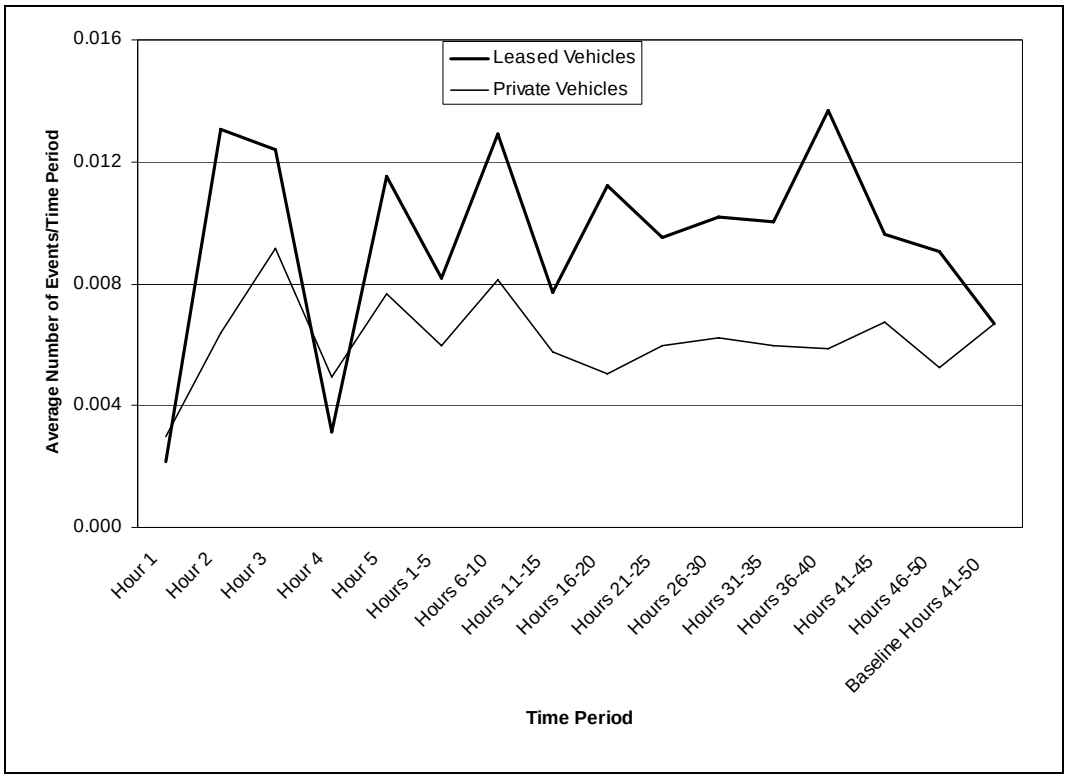


Figure 8.39. Mean number of events per mile for leased and private vehicles for time periods of interest over hours 1-50 of the study.

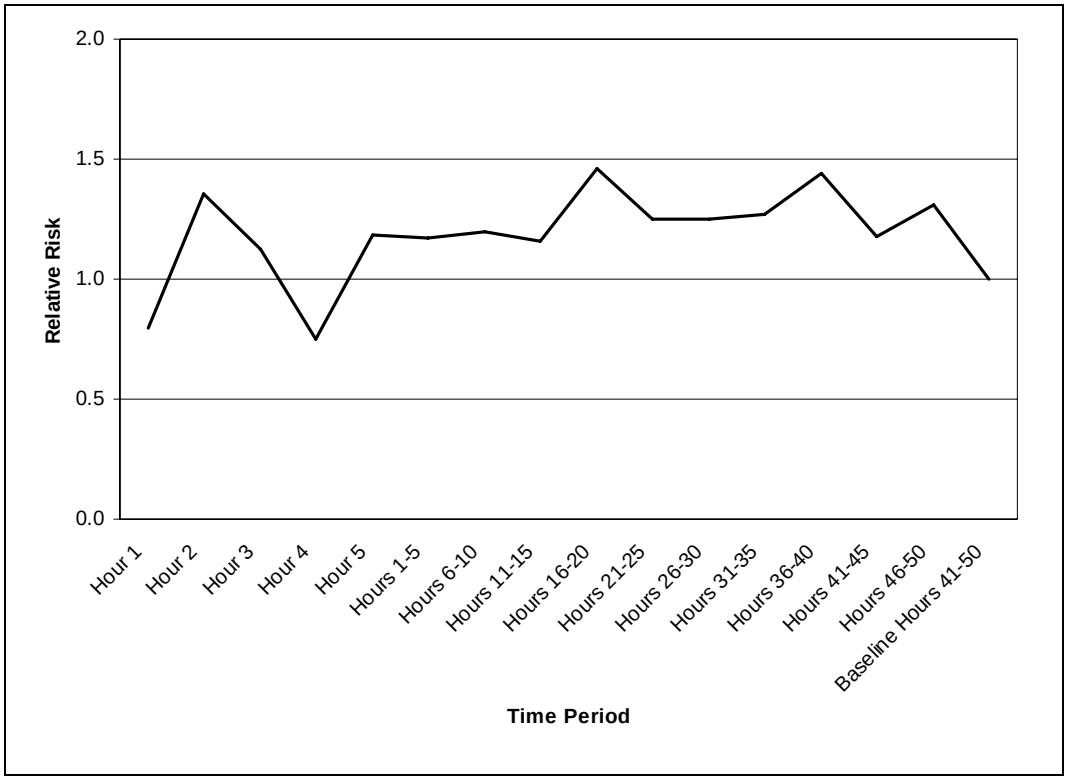


Figure 8.40. Events per mile leased versus private vehicle relative risk for time periods of interest over hours 1-50 of the study.

Question 6. Based on the number and type of valid events per mile driven, is there a significant difference in the relative risk of driving for the same driver for weeks 1 through 4 of driving a privately owned vehicle and weeks 1 through 4 of driving a leased vehicle?

The purpose of this question was to investigate the driver adaptation process for the same driver in a leased vehicle versus a privately owned vehicle (both instrumented) on a per mile basis so that any exposure differences between leased vehicle driving and private vehicle driving could be taken into account. Table 8.19 presents the data available for these analyses in terms of the mean number of events per mile and the RR for each of weeks 1 through 4. Only switch drivers for whom matched data were available for each week were used. For example, if there were no data for Driver 405 for Week 2 of leased vehicle driving, then the private vehicle driving for Driver 405 for Week 2 was also discarded. This resulted in a perfectly matched set of drivers for each week. Weeks 2 through 4 were used as baseline for the RR calculations, and the leased average was set equal to the baseline average to provide for a control condition as was done in previous analyses. Figure 8.41 shows the mean number of events for leased and private vehicle driving for weeks 1-4, while Figure 8.42 presents the RR for these time periods.

As shown in Figure 8.41, there were a greater mean number of events per mile for leased vehicle driving for three of the four weeks, even with a perfectly matched set of leased vehicle drivers. The results of the per mile analysis were in close agreement with Question 3 for younger and

older drivers, so those results are not shown here. The most important finding of this analysis is that even when the same driver was switched from a private vehicle to a leased vehicle, there were still more events per mile in the leased vehicle than in the private vehicle. If the increased number of events in leased vehicle driving for the same driver was due to vehicle unfamiliarity, this effect was not extinguished over the first four weeks, even when exposure was taken into account.

Table 8.19. Events per mile statistics for a matched set of all switch drivers for weeks 1-4.

Week	Leased Events/Mile	Private Events/Mile	RR
Week 1	0.0062	0.0045	1.13
Week 2	0.0083	0.0039	1.33
Week 3	0.0037	0.0031	1.09
Week 4	0.0071	0.0030	1.44
Weeks 2-4 Baseline	0.0033	0.0033	1.00

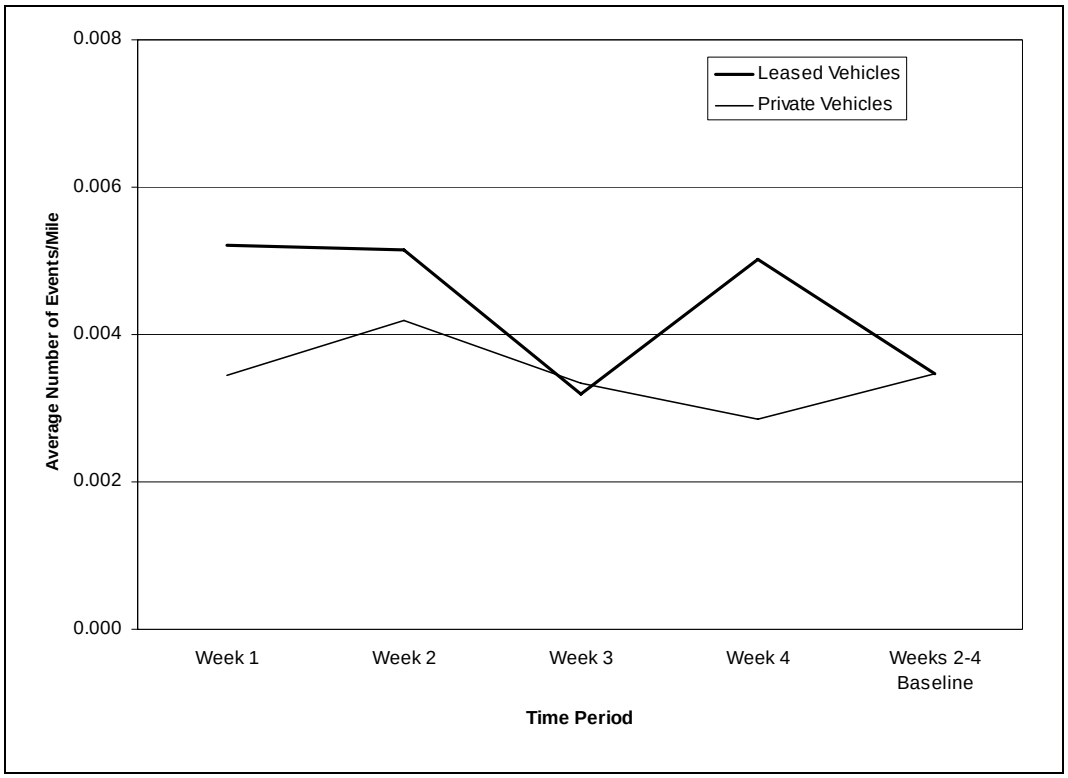


Figure 8.41. Matched set of switch drivers leased versus private vehicle mean events per mile for weeks 1-4.

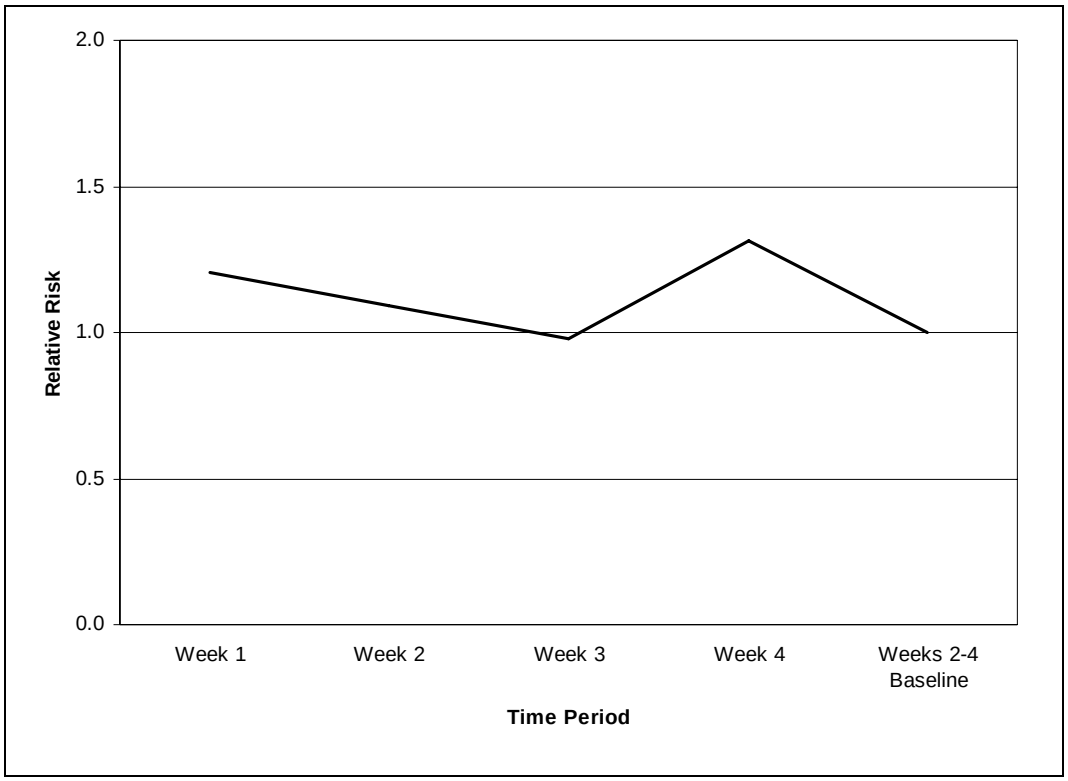


Figure 8.42. Matched set of switch drivers leased versus private vehicle relative risk per mile for weeks 1-4.

The analysis to date for switch drivers has examined only weekly data. However, it was also possible to look at switch drivers in the first hours of driving each vehicle type, using the hourly dataset prepared for Questions 2 and 4. An examination of the hourly data provided some evidence to support the theory that individual younger drivers may have had trouble adapting to a new vehicle. The switch driver data were examined to find drivers for whom 10 hours of data were available for both the private vehicle portion of their driving and the leased vehicle portion. Six of the switch drivers met this criterion. Altogether, these drivers experienced 9 events in the first 10 hours of driving their own private vehicle, and 18 events in the first 10 hours after switching over to a leased vehicle (Figure 8.43). Note that there were no events in the first hour of driving for either leased or private vehicle driving.

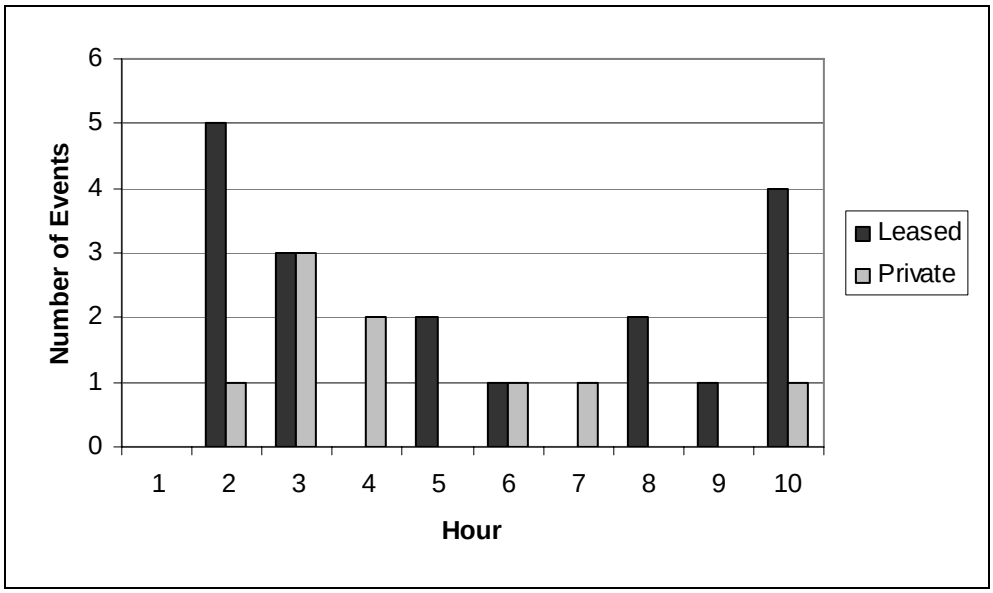


Figure 8.43. Number of events for leased and private vehicle driving in the first 10 hours of driving for a matched set of 6 switch drivers.

When the data was examined in detail, it was discovered that the two older switch drivers experienced 3 events in their private vehicles and none in the leased vehicles, while the 4 younger drivers experienced 6 events in the private vehicles and 18 in the leased vehicle. Figure 8.44 shows the number of events for the 4 younger drivers. Of the 6 events in the private vehicles, there were no near-crashes and 1 crash (a younger driver). Of the 18 events in the leased vehicles, there were 2 near-crashes (both younger drivers) and no crashes. The same pattern held true when events were corrected for exposure (events per mile). Figure 8.45 shows the events per mile for the set of 6 drivers, while Figure 8.46 shows the same thing for the 4 younger drivers only.

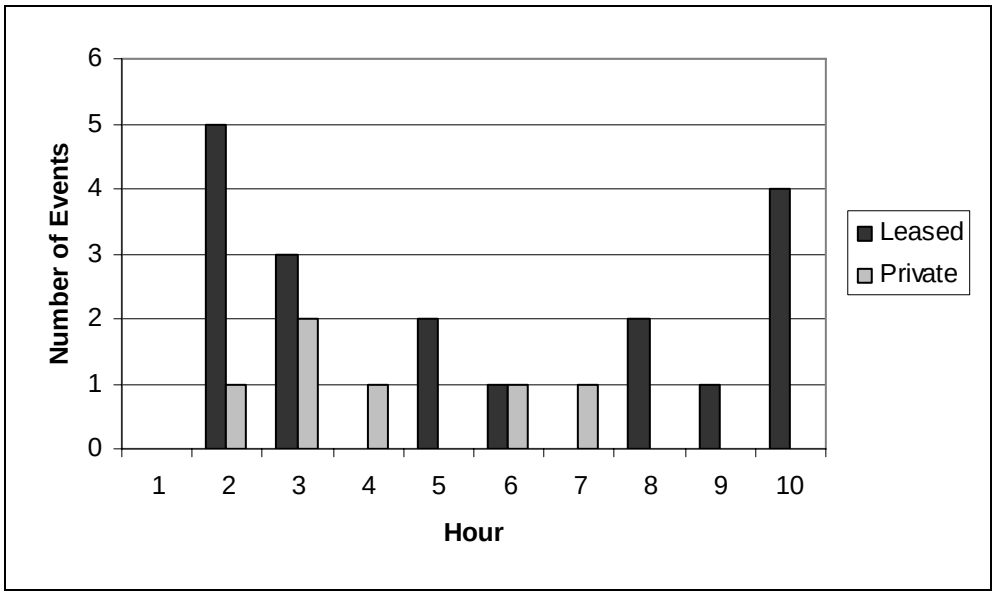


Figure 8.44. Number of events for leased and private vehicle driving in the first 10 hours of driving for a matched set of 4 younger switch drivers.

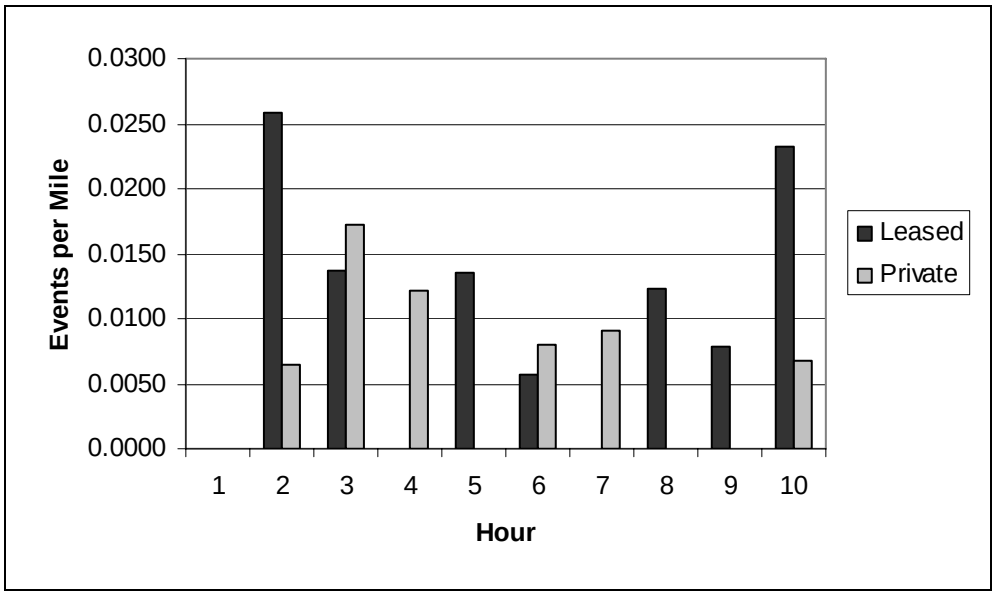


Figure 8.45. Events per mile for leased and private vehicle driving in the first 10 hours of driving for a matched set of 6 switch drivers.

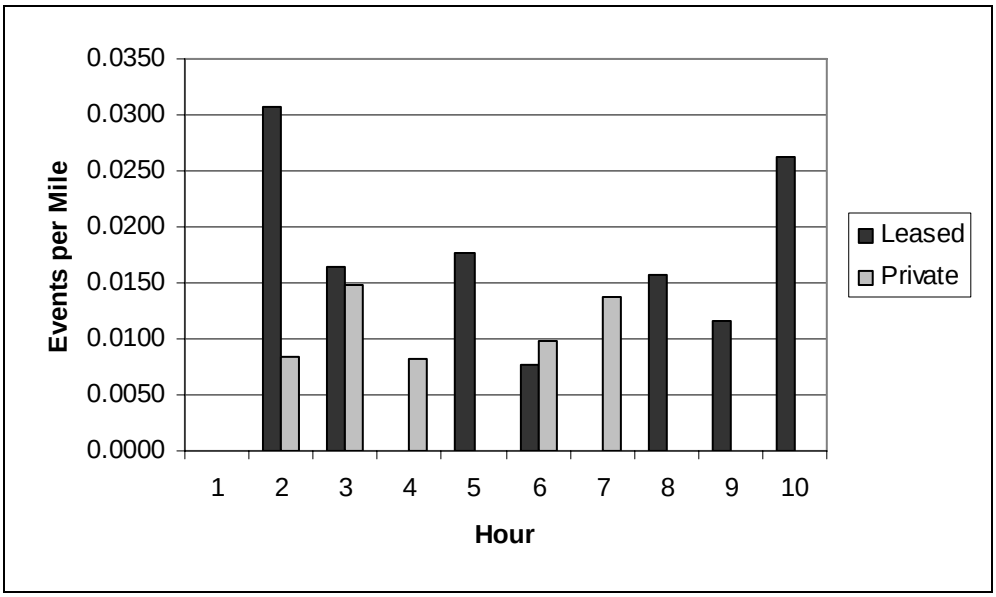


Figure 8.46. Events per mile for leased and private vehicle driving in the first 10 hours of driving for a matched set of 4 younger switch drivers.

The relative risk was then calculated for this rather small dataset using an average of weeks 1-3 as the exposure period and weeks 4-10 as the baseline. The events per mile RR calculated using the same methods described previously is 1.14 for all drivers for weeks 1-3 as compared to weeks 4-10, and 1.20 for the younger drivers for the same time periods.

DISCUSSION

The questions addressed in Chapter 8, *Goal* 4 were intended to explore issues of whether driver behavior in an instrumented vehicle changed over time. The units of time used were weeks (weeks 1 through 50) and hours (the first 50 hours). The issues explored were: (1) driver behavior in a newly instrumented leased vehicle in the first weeks as compared to by the end of the study; (2) driver behavior in the first few hours of driving; and (3) driver behavior for the same driver in four weeks of leased vehicle driving versus four weeks of private vehicle driving. The relative risk analysis technique was borrowed from the field of epidemiology, and required that there be both an exposed and unexposed condition and a comparison and baseline time period. The relative risk of an event for the exposed condition for a given time period can then be calculated as compared to the unexposed condition for that time period. For these questions, the exposed condition was the leased vehicle and the unexposed condition was the private vehicle, since the private vehicle drivers kept driving their usual vehicles while the leased vehicle drivers were exposed to a new vehicle. The baseline time periods were an average of weeks 41-50 for the yearly comparison, hours 41-50 for the hourly comparison, and weeks 2-4 for the leased versus private vehicle comparisons.

In answering Question 1 for driver behavior over the first year of the study, a potential confound between leased vehicles and age was discovered. None of the leased vehicle drivers were over the age of 30, so any results from the leased vehicle analyses may have been confounded by age. An age analysis confirmed that there was indeed an effect of age, but that it was not as large as

the effect for leased versus private vehicles. The next analysis examined whether there were enough younger private vehicle drivers to make a comparison of leased versus private vehicle driving using a matched set of younger drivers. There were approximately 25 younger drivers of both leased and private vehicles. The age distributions of these two groups were quite similar, so these approximately 50 drivers were used for Questions 1, 3, 4, and 6. For Questions 2 and 5, matched sets of switch drivers were used (those who moved from a private vehicle to a leased vehicle at the end of the study).

Question 1 and Question 4: Driving behavior over the year

The purpose of these questions was to investigate the driver adaptation process for leased vehicles and privately owned vehicles with instrumentation over the course of the study. It was expected that drivers would be most adapted to their vehicles and to the instrumentation by the end of the study, so weeks 41-50 were used as the baseline time period. Question 1 explored this issue in terms of the average number of events per vehicle for private and leased vehicles (thus controlling for the different numbers of private and leased vehicles), while Question 4 used events per mile to account for exposure differences in terms of miles driven.

As discussed, a matched set of younger drivers was used to explore these questions. When events were examined, it became obvious that although there was not any appreciable change in driving behavior over the course of a year, there was a consistently higher risk for leased vehicle drivers as compared to private vehicle drivers, even when using a matched set of younger drivers. It was then hypothesized that leased vehicle drivers may be willing to take risks with these vehicles, since they were not responsible for insurance or repairs, and had no ownership interest in the vehicles. If this were the case, it was then hypothesized that the same difference might not occur for crashes and near-crashes, because drivers might be more willing to put a vehicle in harm's way (indicated by a higher number of incidents) but not willing to themselves at risk (as exemplified in the number of crashes and near-crashes). An examination of the near-crash data did not support this hypothesis -- the leased vehicles were still at a higher risk of near-crashes. It was only when the crash data were examined that the hypothesis was supported. The relative risk of crashes over time for this matched set of drivers was relatively low when leased vehicles were compared to private vehicles. Question 4 explored the same issues, except that the data were corrected for exposure in terms of miles driven. The same basic findings appeared when the data were examined in this way. Based on these results, we might expect that if we were to transfer the leased vehicle drivers into their own private vehicles in which they would be responsible for repairs, insurance, etc., that their incident levels would drop to the same levels shown for private vehicle drivers.

Question 2 and Question 5: Driving behavior over the first 50 hours

Questions 2 and 5 were designed to determine whether drivers experienced an increase in valid events over the first few hours of driving a newly instrumented vehicle. The hypothesis was that drivers would not act "naturally" when they first began using an instrumented vehicle, and that they would begin to act more naturally as time went on and they would forget about the cameras and computers. It was hypothesized that the drivers would drive more carefully and experience fewer events when they were aware of the cameras, and that they would revert to normal behavior as time went on. If a point in time can be identified at which drivers adapted and began acting more naturally, this would be useful information for future instrumented vehicle studies of

naturalistic driving. Previous experience at VTTI has indicated that drivers adapt amazingly quickly to the instrumented vehicle (perhaps within minutes, even in an unfamiliar vehicle), but the question has never been empirically analyzed as was attempted here. As seen in Questions 1 and 4, we might expect to find differences in events between leased and private vehicles, even with a matched set of younger drivers, so these two groups of drivers were explored as they were in Questions 1 and 4.

As before, even when controlling for age to the degree possible, leased vehicles experienced a greater mean number of events for nearly every time period studied. The only exceptions were hours 1 and 4, in which the leased and private vehicles experienced nearly identical mean numbers of events. It did appear that drivers of both vehicle types were being very careful during the first hour with a newly instrumented vehicle. The age analysis had the same trend. Both younger and older drivers experienced their lowest mean rate of events in the first hour of driving, and both then leveled off to about the same levels seen in the weekly analysis.

An analysis of near-crashes was also performed. There were 45 leased vehicle near-crashes and 22 private vehicle near-crashes for younger drivers in the first 50 hours of driving. The pattern was the same one seen in the weekly analysis, although there did seem to be an overall decline in near-crashes for leased vehicle drivers over the first 50 hours of driving, while the private vehicle driver near-crash levels remained nearly flat over the 50 hours. For nearly every time period, however, there were a greater mean number of near-crashes for the leased vehicle drivers as compared to the private vehicle drivers, even in this matched set of younger drivers. The results for both overall events and near-crashes were quite similar when controlling for exposure in the per mile analyses in Question 5.
Questions 2 and 5 provided support for the thesis that drivers are more careful when first using an instrumented vehicle. The effect appears to wear off after the first hour. The dataset did not provide a breakdown by minutes, so it was not possible to tell whether this occurred within the first 5 minutes, the first half hour, or whether the adaptation did indeed happen after exactly one hour.

Question 3 and Question 6: Same driver for four weeks in private and leased vehicles

The purpose of questions 3 and 6 was to investigate the driver adaptation process to an unfamiliar vehicle for the same driver in a leased vehicle versus a privately owned vehicle (both instrumented). The analyses for Questions 1, 2, 4, and 5 made it apparent that there was a real increased risk for drivers in leased vehicles as compared to private vehicles, so Questions 3 and 6 were also seen as being useful in further examining this risk. Only switch drivers for whom matched data were available for each week were used. For example, if there were no data for Driver 405 for Week 2 of leased vehicle driving, then the private vehicle driving for Driver 405 for Week 2 was also discarded. This resulted in a perfectly matched set of drivers for each week.

The data did not indicate any clear trend of adaptation to a new vehicle. When examining the leased versus private vehicle question, however, the analyses using perfectly matched sets of switch drivers had similar results to the previous analyses. Even when the same driver was switched from a private vehicle to a leased vehicle, there were still more events in the leased vehicle than in the private vehicle. There was some indication that older switch drivers might not fit this profile for weeks 1 through 3, but by week 4, the difference in the mean number of

events between leased and private vehicles was back at the usual level. If the increased number of events in leased vehicle driving for the same driver was due to vehicle unfamiliarity, this effect was not extinguished over the first four weeks, and based on the yearly results, the higher numbers for leased vehicles likely had very little to do with adaptation, since after 50 weeks there were still clearly more vents and more near-crashes for leased vehicles as compared to private vehicles. The results of the per mile analysis for Question 6 were in close agreement with Question 3 for younger and older drivers. The most important finding from these analysis was that even when the same driver was switched from a private vehicle to a leased vehicle, there were still more events per mile in the leased vehicle than in the private vehicle. If there was an effect of adaptation, it was extinguished before the first week was out.

In order to further explore the issue of adaptation to a new vehicle, the hourly dataset prepared for Questions 2 and 4 was examined over the first 10 hours (representing approximately one weeks' worth of data). The switch driver data were examined to find drivers for whom 10 hours of data were available for both the private vehicle portion of their driving and the leased vehicle portion. Six of the switch drivers met this criterion. There were no events in the first hour of driving for either leased or private vehicle driving for these 6 drivers, providing support for drivers being more careful during the first hour after beginning to drive an instrumented vehicle, but not providing real support for adaptation to a new vehicle.

Altogether, these 6 drivers experienced 9 events in the first 10 hours of driving their own private vehicle and 18 events in the first 10 hours after switching over to a leased vehicle. When only the 4 younger drivers were considered, there were 6 events in the first 10 hours for private vehicles and 18 events for leased vehicles. This trend held true even when the per mile analysis was conducted to control for exposure. This provided some evidence to support the theory that individual younger drivers may have trouble adapting to a new vehicle, while there was no evidence that this was true for older drivers.

CHAPTER 9: GOAL 5, DETERMINE REAR-END CRASH CONTRIBUTING FACTORS AND DYNAMIC CONDITIONS

DATA ANALYSIS OVERVIEW

For the analyses associated with Chapter 9, *Goal* 5, four lead-vehicle conditions were originally slated for evaluation:

- Lead vehicle stopped less than or equal to 2 seconds
- Lead vehicle stopped greater than 2 seconds
- Lead vehicle decelerating
- Lead vehicle moving at a slower constant speed

Upon examination of the events in the 100-Car Study reduced data, a fifth lead-vehicle condition was added:

- Lead vehicle accelerating

Since the vehicles were instrumented with a rear-facing radar system, following-vehicle events (for which the subject vehicle was the lead vehicle) are also included in these analyses. For this chapter, we will examine the frequency and rate of lead and following-vehicle events, then the relation of the events to driver characteristics, contributing factors, and drivers' corrective actions.

Data Included in the Analyses

An important note for these analyses is that, since relative rates of events between groups were of interest, the rate of events was initially calculated at an individual level, and then averages were calculated for the group level. This is in contrast to analyses conducted for Chapter 12, *Goal 8*, which calculated rates based upon total mileage collected in the study.

Therefore, to arrive at the number of million vehicle miles traveled (MVMT) for the rate data, estimates were calculated based upon video reduction during which reductionists viewed a sample of 100 trip files for each vehicle and recorded whether the primary driver was behind the wheel. The percentage of trip files that each of 109 primary drivers were behind the wheel was multiplied by the total vehicle miles traveled for that vehicle (based on the odometer readings) to arrive at a VMT for each primary driver. This was then multiplied by 1,000,000 to provide a MVMT estimate for each driver. Each rate analysis went back to the driver level to determine the rate of events per MVMT based on the drivers involved in those events. Drivers who were not included in a certain cell (with none of the event type shown in that cell) were assigned a zero for MVMT.

From the dataset of 109 primary drivers, 5 had traveled less than 1,000 miles and 9 did not have any lead or following-vehicle conflicts, so they were eliminated from consideration for these analyses. This left a total of 95 drivers in these analyses.

Question 1. What are the relative frequencies of these 5 RE scenarios, and the rates of driver involvement in these scenarios per MVMT?

The frequency of lead-vehicle and following-vehicle events by level of severity was determined for the driver data included in the analyses (Figure 9.1). For the lead-vehicle conflict case, the resulting dataset contained 13 crash events, 268 near-crash events, and 4,747 incidents. For the following-vehicle conflict case, the resulting dataset contained three crash events, 10 near-crash events, and 130 incidents.

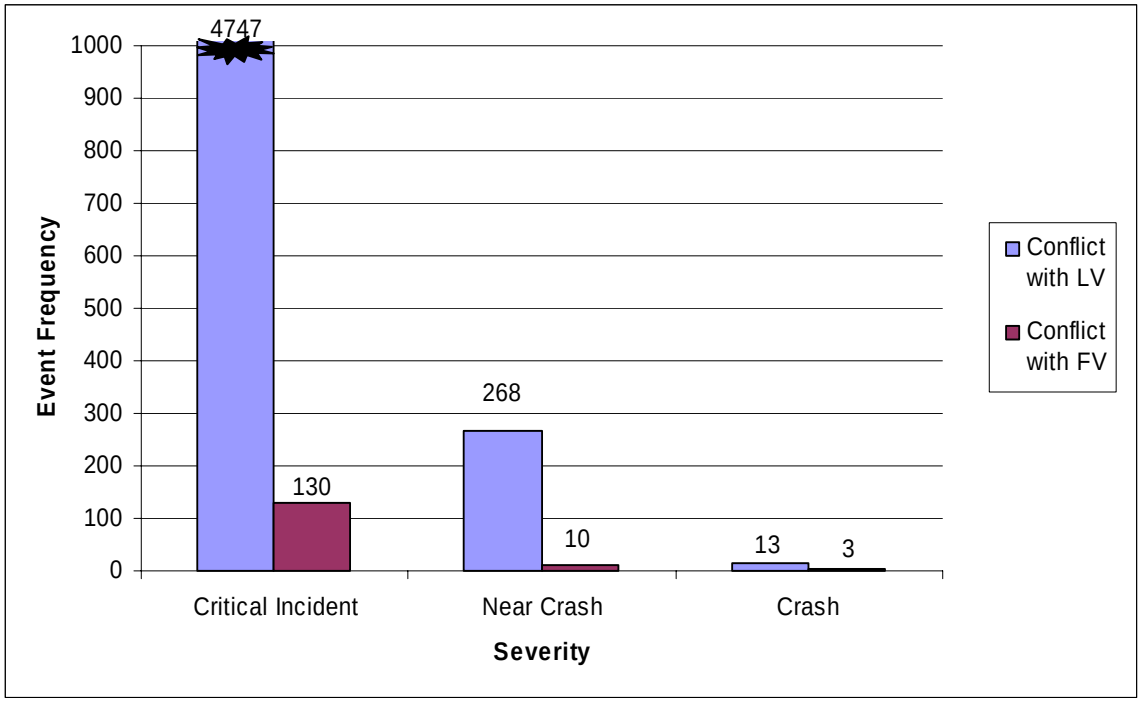


Figure 9.1. Frequency of lead- and following-vehicle events by level of severity.

Note that there were fewer following-vehicle events compared to lead-vehicle events in this dataset. This was due to the differences in the radar signatures for a forward versus a rear-facing radar system. Essentially, a forward-facing radar system has more objects to discern since gaining range on any static object indicates a potential threat. Alternatively, a rear-facing radar system only needs to produce a signature for objects moving toward the vehicle since all other targets are increasing in range as the vehicle moves forward.

Additionally, it was easier to validate triggers for a lead-vehicle scenario versus a following-vehicle scenario. For lead-vehicle conflicts, the radar signatures gave reductionists better data for the rate of deceleration, forward time-to-collision (TTC), and forward range, which could be verified readily using the subject vehicle accelerometer and the forward camera. However, the rear radar did not supply a direct measure of rate of deceleration or speed. For following-vehicle conflicts, the rate of deceleration was much harder to calculate and more difficult to assess by the reductionists with the rear-facing camera. Therefore, verifying conflicts with following vehicles was a more difficult process and only the most severe events were likely to be validated. Based

on the event data available for lead and following vehicles, the event rate per MVMT was calculated for each level of severity (Figure 9.2). Data for Lead-vehicle and Following-Vehicle will be discussed separately in this section.

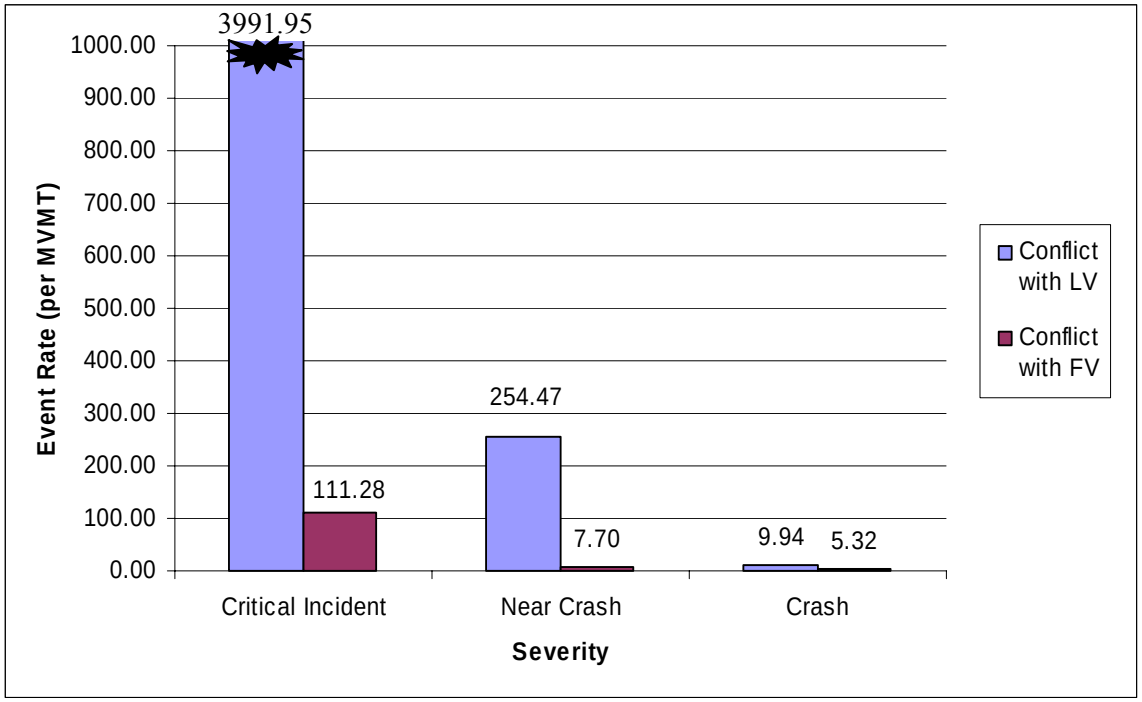


Figure 9.2. The rate of lead- and following-vehicle conflicts by level of severity.

Lead-vehicle (LV) Data

The frequency of lead-vehicle events for each of 5 lead-vehicle scenarios is shown in Table 9.1. It can be seen that the most common scenario for incidents was *LV decelerating*, followed by *LV stopped >2 s*. For near-crashes, the most common scenario was again *LV decelerating*, followed this time by *LV stopped ≤2 s*. It is noteworthy that although *LV decelerating* was the most common scenario for incidents and near-crashes, there were no crashes for this scenario. All of the crashes occurred in circumstances for which the LV was stopped when the crash occurred, either more than 2 seconds (6 crashes) or 2 seconds or less (7 crashes). This means that none of the lead vehicles were still moving at the time the subject vehicle struck them. There were a fairly small number of incidents and near-crashes for LVs moving at a slower, constant speed. There were only 8 incidents for LV accelerating, and no crashes or near-crashes for this scenario.

Table 9.1. Frequencies for the 5 RE lead-vehicle scenarios by event severity

Severity	LV accelerating	LV moving slower, constant speed	LV decelerating	LV stopped ≤ 2 s	LV Stopped > 2 s
Incident	8	119	2,436	989	1,195
Near-Crash	0	5	148	74	41
Crash	0	0	0	7	6

As can be seen in Table 9.2, the overall data pattern remained the same, even when controlled for exposure by calculating rates per MVMT. The rate calculation resulted in no change in the most common scenarios. Tables from this point on will be presented in rate per MVMT format.

Table 9.2. Rate of events per MVMT for the 5 RE lead-vehicle scenarios by event severity

Severity	LV accelerating	LV moving slower, constant speed	LV decelerating	LV stopped ≤ 2 s	LV Stopped > 2 s
Incident	7.35	83.10	2,011.38	838.17	1,051.96
Near-Crash	0	3.68	144.75	67.24	38.80
Crash	0	0	0	5.49	4.45

The rate of events per MVMT by age grouping and severity for each of the 5 scenarios is shown in Table 9.3. The only distinct trend in the data was that 18- to 20-year-olds had the greatest rate of incidents and near-crashes for each of the 5 scenarios. The rates were fairly constant across each of the other age categories for each scenario type for incidents and near-crashes. The 18- to 20-year-olds did have the highest rate of crashes for *LV stopped* $\leq$ *2 s*. For *LV stopped > 2 seconds* crashes, 35-44-year-olds had the highest rate per MVMT.

Table 9.3. Rate of events per MVMT for the 5 RE lead-vehicle scenarios for each age group by event severity

Incident

Subject Age	LV accelerating	LV moving slower, constant speed	LV decelerating	LV stopped ≤ 2 s	LV Stopped > 2 s
18-20	13.11	195.52	4,674.44	1,814.74	2,621.40
21-24	9.82	77.96	1,740.10	679.49	1,017.61
25-34	10.27	48.96	1,315.16	732.51	694.25
35-44	5.67	92.98	2,167.65	672.16	918.96
45-54	4.04	54.41	1,330.93	860.34	741.80
55+	0	33.48	1,019.87	310.58	406.21

Near-Crash

Subject Age	LV accelerating	LV moving slower, constant speed	LV decelerating	LV stopped ≤ 2 s	LV Stopped > 2 s
18-20	0	10.21	471.38	154.62	128.92
21-24	0	0	137.48	36.91	28.33
25-34	0	0	58.66	70.03	34.36
35-44	0	6.15	125.94	78.55	17.23
45-54	0	6.04	56.39	42.46	20.43
55+	0	0	45.97	26.54	13.32

Crash

Subject Age	LV accelerating	LV moving slower, constant speed	LV decelerating	LV stopped ≤ 2 s	LV Stopped > 2 s
18-20	0	0	0	15.07	5.10
21-24	0	0	0	2.78	2.78
25-34	0	0	0	8.41	6.43
35-44	0	0	0	0	11.14
45-54	0	0	0	3.64	0
55+	0	0	0	4.97	0

The next set of analyses provided a further breakdown of rate of RE lead-vehicle scenarios by both age and gender. Age groups were combined into three groups for these analyses (ages 18-24, 25-44, and 45+). The first analysis examined the rate of incidents for the 5 RE lead-vehicle scenarios by age and gender as shown in Table 9.4. For *LV decelerating* (the most common category in terms of frequency), females experienced a higher rate of incidents than males in the all three age groups. For *LV stopped ≤ 2 seconds* and *LV stopped > 2 s*, females had a higher rate in the 18-24 and 45+ categories, while males had a higher rate in the 25- to 44 category. In the LV moving at slower constant speed category, males had a higher incident rate than females for all three categories. The age and gender rates for near-crashes from Table 9.4 were examined next. Females had a higher rate of near-crashes than males in the all three age groups for the three most common scenarios: *LV decelerating*, *LV stopped ≤ 2 s,* and *LV stopped > 2 s*. For crash rates, age and gender comparison were only possible for two groupings, both in the LV

decelerating scenario. The crash rate for 18- to 24-year-old females exceeded that of males by 3.5 to 1, while 25- to 44-year-old females exceeded the male rate by 7.2 to 1.

Table 9.4. Rate of events per MVMT for the 5 RE lead-vehicle scenarios by severity, age group, and gender

Incident

Subject Age	Subject Gender	LV accelerating	LV moving slower, constant speed	LV decelerating	LV stopped ≤ 2 s	LV Stopped > 2 s
18-24	Female	6.54	99.45	3,668.34	1,476.74	1,937.92
	Male	16.78	163.32	2,233.31	8,10.01	1,444.34
25-44	Female	22.51	41.21	2,097.59	631.62	525.74
	Male	2.17	83.45	1,616.70	729.03	920.95
45+	Female	0.00	27.62	1,429.87	935.84	679.49
	Male	3.66	55.76	1,067.02	450.70	553.68

Near-Crash

Subject Age	Subject Gender	LV accelerating	LV moving slower, constant speed	LV decelerating	LV stopped ≤ 2 s	LV Stopped > 2 s
18-24	Female	0.00	0.00	299.01	92.60	86.68
	Male	0.00	9.61	261.70	82.17	54.35
25-44	Female	0.00	0.00	198.98	101.74	32.81
	Male	0.00	4.39	51.59	63.66	22.73
45+	Female	0.00	0.00	62.20	58.37	26.58
	Male	0.00	5.47	46.13	22.76	12.17

Crash

Subject Age	Subject Gender	LV accelerating	LV moving slower, constant speed	LV decelerating	LV stopped ≤ 2 s	LV Stopped > 2 s
18-24	Female	0.00	0.00	0.00	12.06	0.00
	Male	0.00	0.00	0.00	3.44	8.24
25-44	Female	0.00	0.00	0.00	10.74	0.00
	Male	0.00	0.00	0.00	1.49	12.32
45+	Female	0.00	0.00	0.00	0.00	0.00
	Male	0.00	0.00	0.00	6.61	0.00

Following Vehicle Data

As was true for the lead-vehicle scenarios, the following-vehicle events were concentrated in the *SV decelerating* scenario. The next most common scenarios of *SV stopped ≤ 2 seconds* and *SV stopped > 2 seconds* were nearly equal to one another in terms of frequency. Recall that for this scenario, the 100-Car Study subject vehicle (SV) was considered to be the lead vehicle and was struck from behind by a following vehicle. Table 9.5 presents the overall number of following-vehicle events. Table 9.6 presents the same information in rate per MVMT, and the pattern is much the same as for the frequency data. As before, the remainder of tables in this section will be presented for event rate per MVMT.

Table 9.5. Frequencies for the 5 RE following-vehicle scenarios by event severity

Severity	SV accelerating	SV moving slower, constant speed	SV decelerating	SV stopped ≤ 2 s	SV Stopped > 2 s
Incident	1	21	207	48	63
Near-Crash	1	0	26	15	0
Crash	0	0	4	2	4

Table 9.6. Rate of events per MVMT for the 5 RE following-vehicle scenarios by event severity

Severity	SV accelerating	SV moving slower, constant speed	SV decelerating	SV stopped ≤ 2 s	SV Stopped > 2 s
Incident	1.04	14.90	184.23	37.32	51.84
Near-Crash	0.50	0.00	22.74	11.12	0.00
Crash	0.00	0.00	5.70	2.00	2.15

The rate of events by age grouping and severity for each of the 5 scenarios is shown in Table 9.7. Since we only have data for the instrumented vehicle, the driver variables in this case refer to the driver of the lead (struck) vehicle. For the *SV decelerating* scenario, there was a clear decreasing trend for incidents with increasing age. The *SV stopped > 2 seconds* scenario, although less common, also exhibited a trend for decreasing incident rates with increasing age, while the other three scenarios did not show a clear trend. There were few data points in the near-crash and crash data, and no obvious trends or patterns were noted. The only data point that pops out is the crash rate for *SV decelerating* for 25- to 34-year-olds, which was more than three times as high as any of the other crash rates for this table.

Table 9.7. Rate of events per MVMT for the 5 RE following-vehicle scenarios for each age group by event severity

Incident

Subject Age	SV accelerating	SV moving slower, constant speed	SV decelerating	SV stopped ≤ 2 s	SV Stopped > 2 s
18-20	0.00	26.45	373.32	46.25	151.04
21-24	0.00	24.13	220.13	34.54	34.63
25-34	0.00	14.13	210.92	61.18	59.61
35-44	5.67	7.66	163.57	48.91	36.30
45-54	0.00	4.39	91.81	13.88	23.00
55+	0.00	13.51	32.98	14.14	15.08

Near-crash

Subject Age	SV accelerating	SV moving slower, constant speed	SV decelerating	SV stopped ≤ 2 s	SV Stopped > 2 s
18-20	0.00	0.00	29.78	21.77	0.00
21-24	0.00	0.00	45.09	5.00	0.00
25-34	0.00	0.00	3.20	12.63	0.00
35-44	2.75	0.00	28.23	12.31	0.00
45-54	0.00	0.00	20.14	0.00	0.00
55+	0.00	0.00	3.38	19.49	0.00

Crash

Subject Age	SV accelerating	SV moving slower, constant speed	SV decelerating	SV stopped ≤ 2 s	SV Stopped > 2 s
18-20	0.00	0.00	0.00	0.00	0.00
21-24	0.00	0.00	4.92	5.14	3.65
25-34	0.00	0.00	23.82	0.00	5.40
35-44	0.00	0.00	0.00	0.00	2.75
45-54	0.00	0.00	3.47	0.00	0.00
55+	0.00	0.00	0.00	7.88	0.00

The next analyses considered the following-vehicle scenarios by both age and gender, presented in Table 9.8. The most meaningful result for the *SV decelerating* scenario was that the 25- to 44-year-old males had over two times the incident rate as female drivers. In contrast, 18- to 24-year-old female drivers had an incident rate that was three times as high as male drivers for the *LV stopped >2 seconds* scenario. When other age and gender comparisons are possible for incidents, there is no obvious pattern. For some age groups, for some scenarios, females exhibited a higher rate, while for others, males exhibited a higher rate. The 45+ female group had higher incident rates than males for every scenario, but the differences were mostly small. The near-crash rate data from Table 9.8 shows that 18- to 24-year-old females had five times the rate as males for the LV decelerating scenario. There were not enough crash data to provide for meaningful age and gender comparisons.

Table 9.8. Rate of events per MVMT for the 5 RE following-vehicle scenarios by severity, age group, and gender

Incident

Subject Age	Subject Gender	SV accelerating	SV moving slower, constant speed	SV decelerating	SV stopped ≤ 2 s	SV Stopped > 2 s
18-24	Female	0.00	26.00	310.47	30.10	124.76
	Male	0.00	24.11	258.02	50.79	38.16
25-44	Female	10.30	13.28	101.88	60.01	30.91
	Male	0.00	9.85	219.93	52.88	54.23
45+	Female	0.00	15.76	80.48	18.33	26.57
	Male	0.00	3.97	59.06	11.51	15.68

Near-Crash

Subject Age	Subject Gender	SV accelerating	SV moving slower, constant speed	SV decelerating	SV stopped ≤ 2 s	SV Stopped > 2 s
18-24	Female	0.00	0.00	60.80	16.79	0.00
	Male	0.00	0.00	12.20	6.91	0.00
25-44	Female	0.00	0.00	10.30	16.27	0.00
	Male	1.96	0.00	18.29	10.97	0.00
45+	Female	0.00	0.00	8.13	0.00	0.00
	Male	0.00	0.00	15.83	12.99	0.00

Crash

Subject Age	Subject Gender	SV accelerating	SV moving slower, constant speed	SV decelerating	SV stopped ≤ 2 s	SV Stopped > 2 s
18-24	Female	0.00	0.00	5.16	5.40	3.83
	Male	0.00	0.00	0.00	0.00	0.00
25-44	Female	0.00	0.00	0.00	0.00	0.00
	Male	0.00	0.00	16.16	0.00	5.63
45+	Female	0.00	0.00	0.00	0.00	0.00
	Male	0.00	0.00	3.14	5.25	0.00

Question 2. What are the kinematic conditions associated with the above 5 RE scenarios?

The kinematic conditions examined for Question 2 include subject vehicle speed at onset of precipitating factor (7 categories) and time headway (four categories) at onset of precipitating factor. These are considered separately for the lead vehicle and following-vehicle scenarios. As for Question 1, the rate of events per MVMT will be presented and analyzed.

Lead-vehicle Data

Recall that the most common LV scenario was *LV decelerating*, with about twice the incident rate as either of the next two most common scenarios (*LV stopped* $\leq$ *2 seconds* and *LV stopped* > *2 s*). For this scenario, the highest incident rates were in the 21-30 and 31-40 mph bins (Table 9.9). For *LV stopped* $\leq$ *2 seconds* the 11-20 and 21-30 mph bins showed the highest incident rates. The *LV stopped* > *2 seconds* scenario exhibited the highest rate with onset speeds of 21-30 and 31-40 mph. These high rates for the moderate speed ranges likely reflect the prevailing speed limits and high traffic density present in the northern Virginia area in which the study was conducted. For near-crashes, the moderate speed ranges of 21-30 and 31-40 also exhibited the highest rates for all three of the most common scenarios. All of the 13 crashes fell into just four scenarios by speed bins. The *LV stopped* $\leq$ *2 seconds* scenario had rate data for the 0-10, 11-20, and 31-40 mph bins, while the *LV stopped* > *2 seconds* scenario had rate data for the 0-10 mph bin. The low vehicle speeds for crashes as opposed to near-crashes may be an indicator that traffic density was extremely high for these crashes, and the driver made some error in closing rate judgment that led to the crash.

Table 9.9. Rate of events per MVMT for the 5 RE lead-vehicle scenarios for each vehicle speed category by event severity

Incident

Subject Vehicle Speed (MPH)	LV accelerating	LV moving slower, constant speed	LV decelerating	LV stopped ≤ 2 s	LV Stopped > 2 s
0-10	1.23	2.19	34.26	68.55	35.95
11-20	0.00	13.85	247.92	246.74	139.55
21-30	0.00	20.91	636.61	330.42	501.84
31-40	2.94	21.78	657.04	159.12	304.33
41-50	2.45	8.76	271.85	25.33	47.78
51-60	0.72	8.05	111.69	3.47	15.48
60+	0.00	7.57	49.57	4.53	6.61

Near-crash

Subject Vehicle Speed (MPH)	LV accelerating	LV moving slower, constant speed	LV decelerating	LV stopped ≤ 2 s	LV Stopped > 2 s
0-10	0.00	0.75	2.74	9.06	6.68
11-20	0.00	0.65	21.21	10.93	5.39
21-30	0.00	0.00	36.68	26.36	8.22
31-40	0.00	0.00	48.62	18.39	14.58
41-50	0.00	0.00	20.60	1.51	3.16
51-60	0.00	1.15	9.14	0.00	0.77
60+	0.00	1.13	5.76	0.00	0.00

Crash

Subject Vehicle Speed (MPH)	LV accelerating	LV moving slower, constant speed	LV decelerating	LV stopped ≤ 2 s	LV Stopped > 2 s
0-10	0.00	0.00	0.00	1.81	3.89
11-20	0.00	0.00	0.00	1.49	0.00
21-30	0.00	0.00	0.00	0.00	0.00
31-40	0.00	0.00	0.00	2.19	0.00
41-50	0.00	0.00	0.00	0.00	0.00
51-60	0.00	0.00	0.00	0.00	0.00
60+	0.00	0.00	0.00	0.00	0.00

Table 9.10 presents the rate data for subject vehicle headway to the lead-vehicle. For the two *LV stopped* scenarios, the rates for headway < 1 second were much lower than the rates for the other headways (by factors of as much as 7 to 1). For near-crashes, the data for headway at precipitating factor onset shows no clear pattern. The rate data were fairly evenly distributed among the four headway categories of < 1 s, 1-1.99 s, 2-2.99 s, and ≥ 3 s. No clear pattern was discernible for the crash data.

Table 9.10. Rate of events per MVMT for the 5 RE lead-vehicle scenarios for each vehicle headway category by event severity.

Incidents

Subject Vehicle Headway (s)	LV accelerating	LV moving slower, constant speed	LV decelerating	LV stopped ≤ 2 s	LV Stopped > 2 s
< 1	0.70	14.50	444.72	94.84	38.27
1-1.99	1.50	21.45	552.29	280.90	245.41
2-2.99	0.00	14.47	273.42	118.08	267.01
≥ 3	2.93	21.19	484.55	234.12	307.04

Near-crashes

Subject Vehicle Headway (s)	LV accelerating	LV moving slower, constant speed	LV decelerating	LV stopped ≤ 2 s	LV Stopped > 2 s
< 1	0.00	0.00	33.74	13.27	3.19
1-1.99	0.00	2.93	23.67	14.97	13.13
2-2.99	0.00	0.00	24.45	4.74	5.11
≥ 3	0.00	0.00	23.90	20.68	10.82

Crashes

Subject Vehicle Headway (s)	LV accelerating	LV moving slower, constant speed	LV decelerating	LV stopped ≤ 2 s	LV Stopped > 2 s
< 1	0.00	0.00	0.00	0.00	0.00
1-1.99	0.00	0.00	0.00	3.23	0.00
2-2.99	0.00	0.00	0.00	1.72	0.96
≥ 3	0.00	0.00	0.00	0.54	2.39

Following Vehicle Data

For the most common following-vehicle scenario of *SV decelerating*, the speed ranges of 11-20 and 21-30 mph had the highest rates of incidents as seen in Table 9.11. This is somewhat lower than was found for the lead-vehicle incidents. For near-crashes, the speed ranges for the *LV decelerating scenario* with the highest rates were 21-30 and 31-40 mph, which may be an indicator that increasing event onset speed results in increased event severity. For the *SV stopped ≤ 2 seconds* scenario, the highest near-crash rates were found at 0-20 mph. The differences between the near-crash rates for these scenarios point out that the following-vehicle driver may have had difficulty in noticing the difference between a decelerating and stopped vehicle, even when the initial speed was lower for the stopped vehicle scenario. The crash data did not show any discernible pattern, given that the few crash events were distributed among so many scenario and speed cells.

Table 9.11. Rate of events per MVMT for the 5 RE following-vehicle scenarios for each vehicle speed category by event severity

Incident

Subject Vehicle Speed (MPH)	SV accelerating	SV moving slower, constant speed	SV decelerating	SV stopped ≤ 2 s	SV Stopped > 2 s
0-10	1.04	1.67	10.99	6.88	14.07
11-20	0.00	0.50	54.23	16.39	12.04
21-30	0.00	0.88	59.38	5.93	13.58
31-40	0.00	2.00	32.93	6.74	6.00
41-50	0.00	2.91	16.28	0.82	2.75
51-60	0.00	4.86	8.80	0.00	0.43
60+	0.00	2.09	0.40	0.56	0.00

Near-crash

Subject Vehicle Speed (MPH)	SV accelerating	SV moving slower, constant speed	SV decelerating	SV stopped ≤ 2 s	SV Stopped > 2 s
0-10	0.00	0.00	0.00	4.67	0.00
11-20	0.00	0.00	1.21	2.51	0.00
21-30	0.50	0.00	5.29	2.32	0.00
31-40	0.00	0.00	5.34	1.20	0.00
41-50	0.00	0.00	4.12	0.43	0.00
51-60	0.00	0.00	4.43	0.00	0.00
60+	0.00	0.00	2.35	0.00	0.00

Crash

Subject Vehicle Speed (MPH)	SV accelerating	SV moving slower, constant speed	SV decelerating	SV stopped ≤ 2 s	SV Stopped > 2 s
0-10	0.00	0.00	0.95	0.00	1.59
11-20	0.00	0.00	0.99	0.00	0.00
21-30	0.00	0.00	0.00	0.00	0.00
31-40	0.00	0.00	0.00	2.00	0.00
41-50	0.00	0.00	0.00	0.00	0.00
51-60	0.00	0.00	3.77	0.00	0.00
60+	0.00	0.00	0.00	0.00	0.00

Table 9.12 presents the rate of following-vehicle events by the headway categories for each following-vehicle scenario. For incidents, the highest rates were observed at less than 2 seconds headway. For near-crashes, the *SV decelerating* scenario had the highest rate of incidents for the < 1 second headway category, as was also true for the lead-vehicle cases. No clear patterns or trends were present for the crashes.

Table 9.12. Rate of events per MVMT for the 5 RE following-vehicle scenarios for each vehicle headway category by event severity

Incidents

Subject Vehicle Headway (sec)	SV accelerating	SV moving slower, constant speed	SV decelerating	SV stopped ≤ 2 s	SV Stopped > 2 s
< 1	1.04	9.05	75.49	10.91	16.70
1-1.99	0.00	2.75	60.01	12.55	17.12
2-2.99	0.00	0.00	16.79	2.33	6.93
> 3	0.00	3.11	30.73	11.54	8.12

Near-crash

Subject Vehicle Headway (sec)	SV accelerating	SV moving slower, constant speed	SV decelerating	SV stopped ≤ 2 s	SV Stopped > 2 s
< 1	0.50	0.00	17.38	5.11	0.00
1-1.99	0.00	0.00	4.11	2.08	0.00
2-2.99	0.00	0.00	0.40	1.63	0.00
> 3	0.00	0.00	0.85	2.30	0.00

Crash

Subject Vehicle Headway (sec)	SV accelerating	SV moving slower, constant speed	SV decelerating	SV stopped ≤ 2 s	SV Stopped > 2 s
< 1	0.00	0.00	5.70	2.00	0.89
1-1.99	0.00	0.00	0.00	0.00	0.00
2-2.99	0.00	0.00	0.00	0.00	0.00
> 3	0.00	0.00	0.00	0.00	0.70

Question 3. What are the contributing and associated factors for RE crashes for the above 5 RE scenarios?

Question 3 examined the contributing and associated factors for RE crashes. As for previous analyses in this chapter, the rate per MVMT data was examined. Two types of contributing factors were considered.

Driver factors considered in these analyses included:

- Driver Physical/Mental Impairment
- Driver 1 Distracted By
- Willful Behavior
- Driver Proficiency
- Roadway Infrastructure
- Driver 1 Vision Obscured by

Environmental and roadway factors considered included:

- Relation to junction
- Traffic control (counted if traffic control present)
- Roadway Alignment (counted if anything other then straight)
- Weather (counted if anything but sunny)
- Surface Condition (counted if anything but dry)
- Light (counted if anything other then day)
- Traffic density (counted if anything other than free flow)

As for the previous questions, lead-vehicle and following-vehicle events will be considered separately.

Lead-vehicle Data

The rate per MVMT data for the lead-vehicle RE driver factors is shown in Table 9.13. Note that more than one driver factor could be selected for each incident, crash, and near-crash. Overall, the driver contributing factor with the highest rate for incidents was driver proficiency. Referring back to the frequency data for this factor, 64 percent of incidents were coded as being related to driver proficiency. This appeared to serve as a catchall category for nearly two-thirds of incidents. Driver distraction had the next highest rate, and 25 percent of incidents were coded with this factor. For the most common category *(LV decelerating)*, the factor with the third highest rating was willful behavior, while for *LV stopped* $\leq$ *2 s* and *LV stopped* > *2 s,* the third highest rating was for driver physical/mental impairment. For near-crashes, driver proficiency also had the highest rate (from the frequency data, 59 percent of near-crashes were coded with this factor). The overall pattern of data for driver contributing factors for near-crashes was very similar to that observed for the incident data. For the crash data (13 crashes) the factor with the highest rate was driver distraction, followed by driver proficiency and driver mental/physical impairment.

Table 9.13. Rate of events per MVMT for the 5 RE lead-vehicle scenarios for each driver contributing factor by event severity

Incidents

Driver Factors	LV accelerating	LV moving slower, constant speed	LV decelerating	LV stopped ≤ 2 s	LV Stopped > 2 s
Driver Physical/Mental Impairment	0.00	4.30	191.42	63.25	97.42
Driver Distracted By	3.90	13.06	542.55	226.62	297.12
Willful Behavior	1.04	40.09	215.51	41.49	74.43
Driver Proficiency	4.35	24.14	1,260.53	542.85	683.68
Roadway Infrastructure	0.00	0.71	32.49	8.52	9.86
Driver Vision Obscured by	0.00	7.00	156.19	51.08	73.11

Near-Crashes

Driver Factors	LV accelerating	LV moving slower, constant speed	LV decelerating	LV stopped ≤ 2 s	LV Stopped > 2 s
Driver Physical/Mental Impairment	0.00	0.75	16.20	11.37	0.86
Driver Distracted By	0.00	1.80	68.06	35.52	19.49
Willful Behavior	0.00	1.88	22.24	5.87	3.29
Driver Proficiency	0.00	1.80	93.62	37.60	25.02
Roadway Infrastructure	0.00	0.00	1.16	0.00	0.55
Driver Vision Obscured by	0.00	0.00	14.64	9.20	5.97

Crashes

Driver Factors	LV accelerating	LV moving slower, constant speed	LV decelerating	LV stopped ≤ 2 s	LV Stopped > 2 s
Driver Physical/Mental Impairment	0.00	0.00	0.00	0.00	1.10
Driver Distracted By	0.00	0.00	0.00	5.49	3.35
Willful Behavior	0.00	0.00	0.00	0.00	0.00
Driver Proficiency	0.00	0.00	0.00	0.00	1.62
Roadway Infrastructure	0.00	0.00	0.00	0.00	0.00
Driver Vision Obscured by	0.00	0.00	0.00	0.00	0.00

Table 9.14 presents the environmental and roadway contributing factor rate data for lead-vehicle RE events. Multiple environmental and roadway contributing factors could be selected for each incident, near-crash, and crash. Traffic density had the highest incident rate by far of any of the environmental and roadway contributing factors for all 5 RE lead-vehicle scenarios. Relation to junction had the next highest rate for all 5 scenarios, followed by traffic control, light, and then weather. The relative rank of rates within each scenario was very consistent. For near-crashes, traffic density again had the highest rate, followed by light and then relation to junction. Again, the relative rank between scenarios was quite consistent. For the 2 scenarios with crashes, all environmental and roadway factors were listed as contributing factors, with very similar rates across the factors and scenarios.

Table 9.14. Rate of events per MVMT for the 5 RE lead-vehicle scenarios for each environmental and roadway contributing factor by event severity

Incidents

Environmental and Roadway Factors	LV accelerating	LV moving slower, constant speed	LV decelerating	LV stopped ≤ 2 s	LV Stopped > 2 s
Relation to Junction	4.81	21.94	696.16	362.22	563.99
Traffic Control	4.11	19.38	435.32	321.03	520.27
Roadway Alignment	2.19	7.62	181.69	60.91	102.01
Weather	0.00	8.03	244.31	102.05	107.46
Surface Condition	0.00	6.19	154.61	84.88	72.19
Light	4.98	17.25	489.33	227.36	261.48
Traffic Density	6.79	73.17	1,861.01	790.56	932.07

Near-crashes

Environmental and Roadway Factors	LV accelerating	LV moving slower, constant speed	LV decelerating	LV stopped ≤ 2 s	LV Stopped > 2 s
Relation to Junction	0.00	0.65	32.35	29.85	11.08
Traffic Control	0.00	0.65	21.40	23.54	9.14
Roadway Alignment	0.00	0.00	22.78	13.17	1.41
Weather	0.00	0.00	35.69	12.88	2.98
Surface Condition	0.00	0.00	23.91	11.28	2.23
Light	0.00	1.88	54.09	29.88	7.88
Traffic Density	0.00	1.80	120.96	55.59	33.77

Crashes

Environmental and Roadway Factors	LV accelerating	LV moving slower, constant speed	LV decelerating	LV stopped ≤ 2 s	LV Stopped > 2 s
Relation to Junction	0.00	0.00	0.00	2.03	2.39
Traffic Control	0.00	0.00	0.00	2.03	2.39
Roadway Alignment	0.00	0.00	0.00	1.11	1.08
Weather	0.00	0.00	0.00	1.49	2.08
Surface Condition	0.00	0.00	0.00	1.49	2.08
Light	0.00	0.00	0.00	0.54	1.29
Traffic Density	0.00	0.00	0.00	4.00	2.08

Following Vehicle Data

Please recall that the *SV decelerating* scenario was the most common scenario for the RE Following Events (Table 9.15). Note also that the driver factors refer to the driver of the subject vehicle (i.e., the struck vehicle). This implies that the struck driver contributed in some way to the event, which is contrary to the common perception of RE crashes (e.g., the striking vehicle is almost always ticketed). However, the video made it clear that drivers who behaved in unexpected ways might indeed contribute to a RE crash in which they were struck. For example, one driver would unexpectedly brake hard (0.5-0.6g) on the freeway approximately 300 feet behind a slower moving lead vehicle.

For SV *decelerating* incidents, the driver factors with the highest rate were driver proficiency, driver distraction, and physical/mental impairment, the same order as for lead-vehicle incidents. Driver vision obscured was also a prominent factor, which may indicate why the SV driver performed an unexpected maneuver. Near-crash rates were consistent with the incident rates for cells with data. For crashes, driver proficiency for the *SV decelerating* scenario was the factor with the lowest rate; the highest rates were observed for willful behavior and roadway infrastructure.

The roadway and infrastructure contributing factor with the highest incident rate was again traffic density for the *SV decelerating* scenario, followed by relation to junction and traffic control (Table 9.16). The same pattern was observed for 2 other scenarios. Traffic density was also the dominating factor for near-crashes for the *SV decelerating* scenario and the *SV stopped ≤ 2 seconds* scenario. There was no clear dominating contributing factor for the rate of crashes for any of the scenarios.

Table 9.15. Rate of events per MVMT for the 5 RE following-vehicle scenarios for each driver contributing factor by event severity

Incident

Driver Factors	SV accelerating	SV moving slower, constant speed	SV decelerating	SV stopped ≤ 2 s	SV Stopped > 2 s
Driver Physical/Mental Impairment	1.04	2.26	16.82	1.08	1.37
Driver 1 Distracted By	0.00	0.56	27.38	2.12	9.00
Willful Behavior	0.00	2.88	9.93	1.84	0.50
Driver Proficiency	1.04	1.09	73.31	8.59	21.88
Roadway Infrastructure	0.00	0.00	4.76	0.00	1.45
Drier 1 Vision Obscured by	0.00	0.76	6.96	4.85	5.73

Near-crash

Driver Factors	SV accelerating	SV moving slower, constant speed	SV decelerating	SV stopped ≤ 2 s	SV Stopped > 2 s
Driver Physical/Mental Impairment	0.00	0.00	2.81	2.03	0.00
Driver 1 Distracted By	0.00	0.00	7.37	3.37	0.00
Willful Behavior	0.50	0.00	1.67	1.08	0.00
Driver Proficiency	0.00	0.00	9.56	4.74	0.00
Roadway Infrastructure	0.00	0.00	0.96	0.00	0.00
Drier 1 Vision Obscured by	0.00	0.00	2.44	0.00	0.00

Crash

Driver Factors	SV accelerating	SV moving slower, constant speed	SV decelerating	SV stopped ≤ 2 s	SV Stopped > 2 s
Driver Physical/Mental Impairment	0.00	0.00	0.00	0.99	0.00
Driver 1 Distracted By	0.00	0.00	1.55	0.00	1.21
Willful Behavior	0.00	0.00	3.77	0.00	0.00
Driver Proficiency	0.00	0.00	0.95	2.00	0.50
Roadway Infrastructure	0.00	0.00	3.77	0.00	0.00
Drier 1 Vision Obscured by	0.00	0.00	0.00	0.00	0.00

Table 9.16. Rate of events per MVMT for the 5 RE following-vehicle scenarios for each environmental and roadway contributing factor by event severity

Incident

Environmental and Roadway Factors	SV accelerating	SV moving slower, constant speed	SV decelerating	SV stopped ≤ 2 s	SV Stopped > 2 s
Relation to Junction	1.04	0.90	66.20	21.54	28.54
Traffic Control	1.04	0.90	57.79	20.85	23.96
Roadway Alignment	0.00	0.00	19.86	3.20	3.58
Weather	1.04	4.09	28.06	3.58	6.35
Surface Condition	1.04	2.11	9.43	1.20	1.52
Light	0.00	2.88	37.14	9.59	4.77
Traffic Density	1.04	13.11	164.26	32.28	41.92

Near-crash

Environmental and Roadway Factors	LV accelerating	LV moving slower, constant speed	LV decelerating	LV stopped ≤ 2 s	LV Stopped > 2 s
Relation to Junction	0.00	0.00	5.15	6.13	0.00
Traffic Control	0.00	0.00	3.94	5.05	0.00
Roadway Alignment	0.00	0.00	3.99	1.41	0.00
Weather	0.50	0.00	6.45	1.76	0.00
Surface Condition	0.00	0.00	3.06	1.76	0.00
Light	0.50	0.00	3.47	1.57	0.00
Traffic Density	0.50	0.00	18.66	10.14	0.00

Crash

Environmental and Roadway Factors	LV accelerating	LV moving slower, constant speed	LV decelerating	LV stopped ≤ 2 s	LV Stopped > 2 s
Relation to Junction	0.00	0.00	1.94	2.00	1.59
Traffic Control	0.00	0.00	1.94	2.00	1.59
Roadway Alignment	0.00	0.00	0.95	1.01	1.59
Weather	0.00	0.00	4.15	0.99	0.38
Surface Condition	0.00	0.00	0.38	0.00	0.38
Light	0.00	0.00	0.00	0.00	1.06
Traffic Density	0.00	0.00	4.76	2.00	1.26

Question 4. What are the corrective actions associated with the 5 RE scenarios above?

The corrective actions (avoidance maneuvers) considered to answer Question 4 included:

- No avoidance maneuver
- Braking (no lockup)
- Braking (lockup)
- Braking (lockup unknown)
- Releasing brakes
- Steered to left
- Steered to right
- Braked and steered to left
- Braked and steered to right
- Accelerated
- Accelerated and steered to left
- Accelerated and steered to right
- Other actions
- Unknown if driver attempted any corrective action

Because there were so many categories of corrective action, the events were grouped together in terms of severity (incidents + near-crashes + crashes). The avoidance maneuver rate data were considered first, followed by an age and gender analysis.

Lead-vehicle Data

Table 9.17 presents the lead-vehicle corrective action rates for all events. For the *LV decelerating* and *stopped* scenarios, the braking (no lockup) dominated the rate data by factors of around 10 to 1. The next highest rates were for braked and steered to right, braked and steered to left, and braking (lockup unknown). When the LV in a RE event was stopped, the SV response almost always involved some sort of braking activity which was rarely accompanied by steering. For *LV decelerating*, steering left and steering right also had fairly high rates, although the overwhelming choice was still braking. For *LV moving at slower constant speed*, a quite different kinematic situation, braking (no lockup) still had the highest rate, but it was nearly equaled by braked and steered to right and no avoidance maneuver. For *LV accelerating*, also quite different kinematically, steered left and braking (no lockup) had the highest rates.

Table 9.17. Rate of events per MVMT for the 5 RE lead-vehicle scenarios for each corrective action category.

Maneuver	LV accelerating	LV moving slower, constant speed	LV decelerating	LV stopped ≤ 2 s	LV Stopped > 2 s
No avoidance maneuver	0.00	110.73	0.00	93.49	80.01
Braking (no lockup)	106.44	148.53	2,257.70	1,073.67	1,337.90
Braking (lockup)	0.00	0.00	120.90	48.97	71.37
Braking (lockup unknown)	0.00	0.00	128.52	127.92	146.29
Releasing brakes	0.00	0.00	54.94	0.00	0.00
Steered Left	113.33	78.13	156.06	100.23	92.08
Steered to right	76.81	89.27	150.47	74.38	119.53
Braked and steered to left	78.97	90.16	248.17	102.55	148.04
Braked and steered to right	0.00	116.58	275.30	152.78	159.22
Accelerated	0.00	0.00	0.00	0.00	0
Accelerated and steered to left	0.00	74.41	136.47	58.20	76.39
Accelerated and steered to right	0.00	62.33	144.86	48.04	64.97
Other action	0.00	0.00	0.00	0.00	55.51
Unknown if driver attempted any corrective action	0.00	0.00	0.00	0.00	81.67

The corrective action rates per MVMT for lead-vehicle RE scenarios are presented in Table 9.18. The rate of braking (no lockup) was higher than for females in every age category. Females also had a higher rate for braked and steered to left for every age category. Older drivers (45+) had a clearly lower rate of braking and steering to the right or left than did the younger age groups. There were no other notable age or gender rate differences.

Table 9.18. Rate of events per MVMT for RE lead-vehicle scenarios by age and gender for each corrective action category.

Maneuver	18-24		25-44		45+	
	M	F	M	F	M	F
No avoidance maneuver	0.00	115.66	70.43	118.18	0.00	0.00
Braking (no lockup)	4,354.69	7,875.62	3,389.60	4,136.37	2,198.70	3,871.41
Braking (lockup)	0.00	71.49	74.11	0.00	62.95	164.74
Braking (lockup unknown)	139.90	169.98	200.38	197.49	100.97	285.71
Releasing brakes	0.00	0.00	54.94	0.00	0.00	0.00
Steered Left	154.86	253.12	151.41	408.77	108.32	106.69
Steered to right	236.73	233.10	169.68	172.41	125.68	82.89
Braked and steered to left	248.68	451.61	238.24	838.92	136.90	181.62
Braked and steered to right	585.53	440.83	358.95	674.58	138.79	200.78
Accelerated	0.00	0.00	0.00	0.00	0.00	0.00
Accelerated and steered to left	163.91	115.31	122.53	0.00	94.06	122.43
Accelerated and steered to right	294.84	82.04	78.54	113.33	64.80	0.00
Other action	0.00	0.00	55.51	0.00	0.00	0.00
Unknown if driver attempted any corrective action	81.67	0.00	0.00	0.00	0.00	0.00

Following Vehicle Data

The corrective action rate data for the following-vehicle scenarios was quite different than that observed for the lead-vehicle scenarios, most likely owing to the fact that it was the lead vehicle's (subject vehicle) response to the threat from behind that was coded. For the most common category of *LV decelerating,* the response with the highest rate was accelerated and steered to right, followed by braked and steered to left, braking (no lockup), and no avoidance maneuver (Table 9.19). Overall, accelerating and no avoidance maneuvers were quite common responses to following-vehicle situations, whereas they were quite uncommon for lead-vehicle situations. Again, this indicates that the SV was either unaware of the threat from behind (no action taken) or that they were aware and were trying to increase the distance between vehicles (acceleration). For cases in which the lead vehicle braked or braked and steered, there may also have been a LV threat that the SV driver perceived as more urgent. The other 3 scenarios with rate data followed the same general pattern.

Table 9.19. Rate of events per MVMT for the 5 RE following-vehicle scenarios for each corrective action category.

Maneuver	SV accelerating	SV moving slower, constant speed	SV decelerating	SV stopped ≤ 2 s	SV Stopped > 2 s
No avoidance maneuver	113.33	105.09	154.46	122.22	114.63
Braking (no lockup)	0.00	82.45	322.71	127.89	164.29
Braking (lockup)	0.00	0.00	89.51	0.00	0.00
Braking (lockup unknown)	0.00	0.00	83.61	0.00	0.00
Releasing brakes	0.00	0.00	0.00	0.00	108.04
Steered Left	0.00	0.00	54.94	0.00	0.00
Steered to right	0.00	0.00	0.00	0.00	0.00
Braked and steered to left	54.94	0.00	222.10	113.33	0.00
Braked and steered to right	0.00	75.51	92.06	60.88	0.00
Accelerated	0.00	68.94	64.84	0.00	74.26
Accelerated and steered to left	0.00	0.00	0.00	59.03	0.00
Accelerated and steered to right	0.00	0.00	410.85	117.91	0.00
Other action	0.00	0.00	0.00	0.00	0.00
Unknown if driver attempted any corrective action	0.00	0.00	0.00	0.00	0.00

There were no obvious age and gender rate patterns observed for following-vehicle events (Table 9.20). For those cells for which age and gender comparisons were possible, the rates were fairly constant across both age and gender. One exception was in the 25- to 44-year-old age group, in which males had a rate that was about twice as high for females for braked and steered left, while the female rate for braked and steered right was nearly twice as high as for males. There is no ready explanation for this finding, however. Another interesting finding was all of the drivers in the 45+ age group exhibited only three responses: braking (no lockup), no avoidance maneuver, and accelerating. Finally, 18- to 24-year-old males had a rate of braked and steered to right that was nearly twice that of females in this age group, while the reverse was true for the braking (no lockup) avoidance maneuver.

Table 9.20. Rate of events per MVMT for RE following-vehicle scenarios by age and gender for each corrective action category.

Maneuver	18-24		25-44		45+	
	M	F	M	F	M	F
No avoidance maneuver	216.49	204.24	239.56	243.01	142.22	131.13
Braking (no lockup)	359.95	639.57	396.77	368.33	142.64	203.73
Braking (lockup)	63.34	0.00	115.67	0.00	0.00	0.00
Braking (lockup unknown)	0.00	57.37	96.73	0.00	0.00	0.00
Releasing brakes	0.00	108.04	0.00	0.00	0.00	0.00
Steered Left	0.00	0.00	54.94	0.00	0.00	0.00
Steered to right	0.00	0.00	0.00	0.00	0.00	0.00
Braked and steered to left	150.88	0.00	190.13	113.33	0.00	0.00
Braked and steered to right	91.16	46.37	60.86	115.75	0.00	0.00
Accelerated	80.32	76.62	54.63	0.00	69.55	0.00
Accelerated and steered to left	59.03	0.00	0.00	0.00	0.00	0.00
Accelerated and steered to right	0.00	0.00	264.38	0.00	0.00	0.00
Other action	0.00	0.00	0.00	0.00	0.00	0.00
Unknown if driver attempted any corrective action	0.00	0.00	0.00	0.00	0.00	0.00

DISCUSSION

The frequency of lead-vehicle and following-vehicle events by level of severity was determined for the driver data included in the analyses. For the lead-vehicle conflict case, the resulting dataset contained 13 crashes, 268 near-crashes, and 4,747 incidents. For the following-vehicle conflict case, the resulting dataset contained 9 crashes, 30 near-crashes, and 239 incidents.

There were fewer following-vehicle events compared to lead-vehicle events in this dataset. This was due to the differences in the radar signatures for a forward versus a rear-facing radar system. Essentially, a forward-facing radar system has many more objects to discern since gaining range on any static object indicates a potential threat. Alternatively, a rear-facing radar system only needs to produce a signature for objects moving toward the vehicle since all other targets are increasing in range as the vehicle moves forward.

Additionally, it was easier to validate triggers for a lead-vehicle scenario versus a following-vehicle scenario. For lead-vehicle conflicts, the radar signatures gave reductionists better data for the rate of deceleration, forward TTC, and forward range, which could be verified readily using the subject vehicle accelerometer and the forward camera. However, the rear radar did not supply a direct measure of rate of deceleration or speed. For following-vehicle conflicts, the rate of deceleration was much harder to calculate and more difficult to assess by the reductionists with the rear-facing camera. Therefore, verifying conflicts with following vehicles was a more difficult process and only the most severe events were likely to be validated.

The four questions answered for this goal addressed driver characteristics, kinematic characteristics, contributing factors, and corrective action for RE events. All data were presented in the form of event rate per million vehicle miles traveled (MVMT). The 5 RE scenarios

considered were *LV accelerating*, *LV moving at slower constant speed*, *LV decelerating*, *LV stopped ≤ 2 s*, and *LV stopped > 2 s*.

Driver characteristics examined the role of age and gender for lead-vehicle and following-vehicle events. The only distinct trend in the lead-vehicle age data was that 18- to 20-year-olds had the highest rate of incidents and near-crashes for each of the 5 scenarios. When age and gender were considered for *LV decelerating* (the most common category in terms of frequency), females experienced a higher rate of incidents than males in the all three age groups. For *LV stopped ≤ 2 seconds* and *LV stopped > 2 s*, females had a higher rate in the 18- to 24 and 45+ categories, while males had a higher rate in the 25- to 44 category. In the *LV moving at slower constant speed* category, males had a higher incident rate than females for all three categories. Near-crashes were examined next. Females had a higher rate of near-crashes than males in the all three age groups for the three most common scenarios: *LV decelerating*, *LV stopped ≤ 2 s,* and *LV stopped > 2 s*. No conclusions could be drawn with regard to crash rates.

As was true for the lead-vehicle scenarios, the following-vehicle events were concentrated in the *SV decelerating* scenario. The next most common scenarios of *SV stopped ≤ 2 seconds* and *SV stopped > 2 seconds* were nearly equal to one another in terms of frequency. For the *SV decelerating* scenario, there was a clear decreasing trend for incidents with increasing age. The *SV stopped > 2 seconds* scenario, although less common, also exhibited a trend for decreasing incident rates with increasing age, while the other three scenarios did not show a clear trend. There were few data points in the near-crash and crash data, and no obvious trends or patterns were noted. The next analyses considered the following-vehicle scenarios by both age and gender. The most meaningful result for the *SV decelerating* scenario was that the 25- to 44-year-old males had over two times the incident rate as female drivers. In contrast, 18- to 24-year-old female drivers had an incident rate that was three times as high as male drivers for the *LV stopped >2 seconds* scenario. When other age and gender comparisons are possible for incidents, there is no obvious pattern. The near-crash rate data shows that 18- to 24-year-old females had five times the rate as males for the LV decelerating scenario.

The next analyses considered the kinematic conditions for RE lead-vehicle events. For lead-vehicle events, incidents and near-crashes had the highest rates for moderate speeds of 21-40 mph, while crashes had the highest rates at lower speeds of 0-20 mph. The high incident and near-crash rates for the moderate speed ranges likely reflect the prevailing speed limits and high traffic density present in the northern Virginia area where the study was conducted. The low vehicle speeds for crashes as opposed to near-crashes may be an indicator that traffic density was extremely high for these crashes, and that the driver made some error in closing rate judgment that led to the crash. For the two *LV stopped* scenarios, the rates for headway < 1 second were much lower than the rates for the other headways (by factors of as much as 7 to 1). For near-crashes, the data for headway at precipitating factor onset shows no clear pattern. The rate data were fairly evenly distributed among the four headway categories of < 1 s, 1-1.99 s, 2-2.99 s, and ≥ 3s. No clear pattern was discernible for the crash data.

For the most common following-vehicle scenario of *SV decelerating*, the speed ranges of 11-20 and 21-30 mph had the highest rates of incidents. This is a somewhat lower speed range than was found for the lead-vehicle incidents. For near-crashes, the speed ranges for the *LV*

decelerating scenario with the highest rates were 21-30 and 31-40 mph, which may be an indicator that increasing event onset speed results in increased event severity. For the *SV stopped ≤ 2 seconds* scenario, the highest near-crash rates were found at 0-20 mph. The differences between the near-crash rates for these scenarios point out that the following-vehicle driver may have had difficulty in noticing the difference between a decelerating and stopped vehicle, even when the initial onset speed was lower for the stopped vehicle scenario. The crash data did not show any discernible pattern, given that the few crash events were distributed among so many scenario and speed cells. For headway, the highest incident rates were observed at less than 2 seconds headway. For near-crashes, the *SV decelerating* scenario had the highest rate of incidents for the < 1 second headway category, as was also true for the lead-vehicle cases. No clear patterns or trends were present for the crashes.

Overall, the driver contributing factor with the highest rate for lead-vehicle incidents was driver proficiency. This appeared to serve as a catchall category for nearly two-thirds of incidents. Driver distraction had the next highest rate. For the most common category *(LV decelerating)*, the factor with the third highest rating was willful behavior, while for *LV stopped ≤ 2 seconds* and *LV stopped > 2 seconds* the third highest rating was for driver physical/mental impairment. For near-crashes, driver proficiency also had the highest rate. The overall pattern of data for driver contributing factors for near-crashes was very similar to that observed for the incident data. For the crash data (13 crashes), the factor with the highest rate was driver distraction, followed by driver proficiency and then driver mental/physical impairment.

For following-vehicle events, the driver factors refer to the driver of the subject vehicle (i.e., the struck vehicle). This implies that the struck driver contributed in some way to the event, which is contrary to the common perception of RE crashes (e.g., the striking vehicle is almost always ticketed). However, the video made it clear that drivers who behaved in unexpected ways might indeed contribute to a RE crash in which they were struck. For example, the video showed drivers who were distracted, did not realize traffic was slowing or stopping, and had to brake hard to avoid a rear-end striking collision. This, in turn, made it more difficult for the following-vehicle driver to stop in time which sometimes led to a rear-end struck collision. For SV *decelerating* incidents, the factors with the highest rate were driver proficiency, driver distraction, and physical/mental impairment, the same order as for lead-vehicle incidents. Driver vision obscured was also a prominent factor, which may indicate why the SV driver performed an unexpected maneuver. Near-crash rates were consistent with the incident rates for cells with data. For crashes, driver proficiency for the *SV decelerating* scenario was the factor with the lowest rate; the highest rates were observed for willful behavior and roadway infrastructure.

The environmental and roadway contributing factors for lead-vehicle RE events was considered next. *Traffic density* had the highest incident rate by far of any of the environmental and roadway contributing factors for all 5 RE lead-vehicle scenarios. *Relation to junction* had the next highest rate for all five scenarios, followed by *traffic control*, *light*, and then *weather*. The relative rank of rates within each scenario was very consistent. For near-crashes, *traffic density* again had the highest rate, followed by *light* and then *relation to junction*. Again, the relative rank between scenarios was quite consistent. The following vehicle roadway and infrastructure contributing factor with the highest incident rate was again traffic density for the *SV decelerating* scenario, followed by relation to junction and traffic control. The same pattern was observed for

two of the other scenarios. Traffic density was also the dominating factor for near-crashes for the *SV decelerating* scenario and the *SV stopped* $\leq$ *2 seconds* scenario. There was no clear dominating contributing factor for the rate of crashes for any of the scenarios.

Corrective actions for lead-vehicle events were considered next. For the lead-vehicle *LV decelerating* and *stopped* scenarios, braking (no lockup) dominated the rate data by factors of around 10 to 1. The next highest rates were for braked and steered to right, braked and steered to left, and braking (lockup unknown). When the LV in a RE event was stopped, the SV response overwhelmingly involved some sort of braking activity, usually without steering. For *LV decelerating*, steering left and steering right also had fairly high rates, although the overwhelming choice was still braking. For *LV moving at slower constant speed*, a quite different kinematic situation, braking (no lockup) still had the highest rate, but it was nearly equaled by braked and steered to right and no avoidance maneuver. For *LV accelerating*, also quite different kinematically, steered left and braking (no lockup) had the highest rates. When age and gender were considered, the rate of braking (no lockup) was higher for females than for males in every age category. Females also had a higher rate for braked and steered to left for every age category. Older drivers (45+) had a clearly lower rate of "braking and steering" (either to the right or left) than did the younger age groups.

The corrective action rate data for the following-vehicle scenarios was quite different than that observed for the lead-vehicle scenarios, most likely owing to the fact that it was the lead vehicle's (subject vehicle) response to the threat from behind that was coded. For the most common category of *LV decelerating,* the response with the highest rate was accelerated and steered to right, followed by braked and steered to left, braking (no lockup), and no avoidance maneuver. Overall, accelerating and no avoidance maneuvers were quite common responses to following-vehicle situations, whereas they were quite uncommon for lead-vehicle situations. Again, this indicates that the SV was either unaware of the threat from behind (no action taken) or that they were aware and were trying to increase the distance between vehicles (acceleration). For cases in which the lead vehicle braked or braked and steered, there may also have been a LV threat that the SV driver perceived as more urgent. The other three scenarios with rate data followed the same general pattern. For the following-vehicle events for which age and gender comparisons were possible, the rates were fairly constant across both age and gender. One exception was in the 25- to 44-year-old age group, in which males had a rate that was about twice as high for females for braked and steered left, while the female rate for braked and steered right was nearly twice as high as for males. Another interesting finding was all of the drivers in the 45+ age group exhibited only three responses: braking (no lockup), no avoidance maneuver, and accelerating. Finally, 18- to 24-year-old males had a rate of braked and steered to right that was nearly twice that of females in this age group, while the reverse pattern was seen for the braking (no lockup) avoidance maneuver.

Additional insight into RE events can be found in Chapter 10, *Goal* 6 and Chapter 11, *Goal* 7. The relationships between the relative frequency of crashes, near-crashes, and incidents for RE events are explored in Chapter 12, *Goal* 8 using Heinrich's Triangles.

This page intentionally left blank.

CHAPTER 10: GOAL 6, DETERMINE LANE CHANGE CONTRIBUTING FACTORS AND DYNAMIC CONDITIONS

BACKGROUND

A primary overall goal for the 100-Car Study was to determine the causes and contributing factors associated with rear-end crashes. An important aspect of fully understanding the rear-end crash problem is understanding the pre-event maneuvers and precipitating factors that, in conjunction with other contributing factors, lead to rear-end crashes.

The purpose of the analyses for Chapter 10, Goal 6 is to understand the degree to which lane change events, such as cut-ins, lead to rear-end conflicts. This has important implications for the design of future forward collision warning systems, since a cut-in vehicle may not provide a radar signature until very late in a conflict scenario.

DATA ANALYSIS OVERVIEW

To begin to understand this issue, the RE conflict data was analyzed for both lead-vehicle (i.e., subject vehicle as following vehicle) and following-vehicle (i.e., subject vehicle as lead-vehicle) scenarios. This research objective analyzed the frequency and rate distributions for the following type of conflicts:

- Rear-end conflicts, both striking and struck.
- Rear-end conflicts resulting from a lane change by other lead-vehicle driver or subject vehicle driver.
- The corrective actions that were taken for all of the above type of scenarios.

Frequency distributions were generated to allow calculation of the rate of occurrence of these types of scenarios per MVMT, the initial kinematic conditions that occurred for each, and the contributing factors that played a role for each type of scenario.

Data Included in the Analyses

Questions 1 and 2 were answered with the same set of data. For the descriptive statistics reported in this chapter, data from 95 drivers were used. We arrived at this dataset by starting with the 109 primary drivers. Of those, 5 had traveled less than 1,000 miles and 9 did not have any lead- or following-vehicle conflicts.

To arrive at the number of MVMT, estimates were calculated based on video reduction during which reductionists viewed a sample of 100 trip files for each vehicle and recorded whether the primary driver was behind the wheel. The percentage of trip files that the primary driver was behind the wheel was multiplied by the total vehicle miles traveled for that vehicle (based on the odometer readings) to arrive at a VMT for each primary driver. The resulting number was then multiplied by 1,000,000 to arrive at MVMT.

For Question 3, data from the side sensors were analyzed. During the course of the 100-Car Study data collection, 20 of the leased vehicles were retrofitted with side sensors. Since more than one driver operated many of the leased vehicles, data from 37 drivers were included. The

side-sensor-equipped vehicles generated 412 weeks of data from the 4,534 total weeks for the study. In addition, sideswipe lane change events were also included to give a complete picture of the lane change conflicts.

Question 1. What was the frequency and rate per MVMT for RE conflicts with lane-change-related initial conditions (same travel lane, LV changed in front of subject vehicle, SV changed in front of following vehicle, SV changed lanes behind LV, following vehicle changes lanes behind SV)?

The breakdown of the frequencies and rates per MVMT of lead- and following-vehicle conflicts for different lane change maneuvers is shown in Tables 10.1 and 10.2. For the lead-vehicle conflicts (i.e., subject vehicle in conflict with a lead vehicle), the vast majority of rear-end events in all levels of severity occurred when no lane change was present. No crashes occurred when there was a lane change as a precipitating factor in front of the subject vehicle or when there was a lane change behind the subject vehicle. There were, however, 64 near-crashes and 324 incidents that occurred when there was a "cut in" to the lane in front of the subject vehicle as compared to only 4 near-crashes and 77 incidents when the SV changed behind a lead vehicle.

As will be described in a later section, the subject vehicle drivers were judged to be impaired (30 events more), distracted (44 events more) and make proficiency-related errors (e.g., inappropriate reaction -- 55 events more) more often in the cases when they were cut-off than when the subject vehicle performed a close lane change in front of another vehicle. This seems to support what has been found throughout this report. At least two elements are required for a conflict to occur and it most often requires a precipitating maneuver plus another contributing factor (often driver state-related). This was primarily true because the drivers were presumably alert and attentive when they were actively performing the lane change maneuver. This logic holds true when looking at the struck vehicle (i.e., subject vehicle conflicts with following vehicles) in Table 10.1, since there were many more cases in which a conflict occurred when the subject vehicle cut-in, creating a rear conflict with the other vehicle.

Table 10.1. Frequencies of conflicts with lead vehicles and conflict with following-vehicle conflicts associated with different lane change maneuvers.

	SV Striking			SV Struck		
	SV lane change behind LV	LV lane change in front of SV	Non-lane change SV/LV conflicts	FV lane change behind SV	SV lane change in front of FV	Non-lane change FV/ SV conflicts
Incident	77	324	4,856	7	307	362
Near-Crash	4	64	280	0	15	44
Crash	0	0	14	0	0	10

Table 10.2. Rates per MVMT of conflicts with lead vehicles and conflicts with following vehicles associated with different lane change maneuvers.

	SV Striking			SV Struck		
	SV lane change behind vehicle	LV lane change in front of SV	Non-lane change SV strikes	FV lane change behind SV	SV lane change in front of FV	Non-lane change struck SV
Incident	65.97	266.79	4,081.80	5.81	241.49	310.22
Near-crash	2.38	54.31	267.49	0.00	13.80	36.16
Crash	0.00	0.00	10.95	0.00	0.00	9.86

An interesting aspect of Tables 10.1 and 10.2 is the difference between the subject vehicle (SV) striking (conflict with lead vehicle) and SV struck (conflict with following vehicle) numbers. As described in Chapter 2, Method, the trigger criteria for the forward radar were easier to filter due to the presence of the on-board accelerometers in the subject vehicle. Therefore, to avoid a large number of false positives toward the rear, the radar-based criteria were set much more stringently. This explains the large overall difference between the left and right sides of Table 10.1. Note, however, the large number of events, compared to the other cells for *SV Struck*, in which the subject vehicle changed lanes in front of another vehicle. All radar and other issues aside, this supports the hypothesis (implied in the previous paragraph) that the most significant issue in lane change maneuvering is a lead-vehicle cut-in when the following driver is required to make a timely reaction.

SV Striking Data

The next analyses examined the rate per MVMT of SV striking events for lane change maneuvers by the driver, characteristic of age. Figure 10.1 provides the rate data for incidents. The no lane change category dominates, as discussed previously, but will be ignored in this discussion since these events are explored more fully in Chapter 9, *Goal* 5. For lane change events, the 18- to 20-year-olds had the highest rate for *LV lane change in front of SV*, while the other age categories had fairly equal rates for this scenario. For the *LV lane change behind SV* scenario, the 21- to 24-year-olds had the highest rate, although the rates were fairly even across age groupings. Near-crashes are considered in Figure 10.2. Again, 18- to 20-year-olds had the highest rate for the *LV lane change in front of SV* scenario by a factor of 2 to 1 over the next highest age group, 35-44-year-olds. The S*V lane change behind LV* scenario had too few cases to show any clear age trends. Since there were no lane change crashes, the crash data will not be shown or discussed in this chapter. The driver characteristics for all crashes and RE crashes are explored further in other chapters of this report.

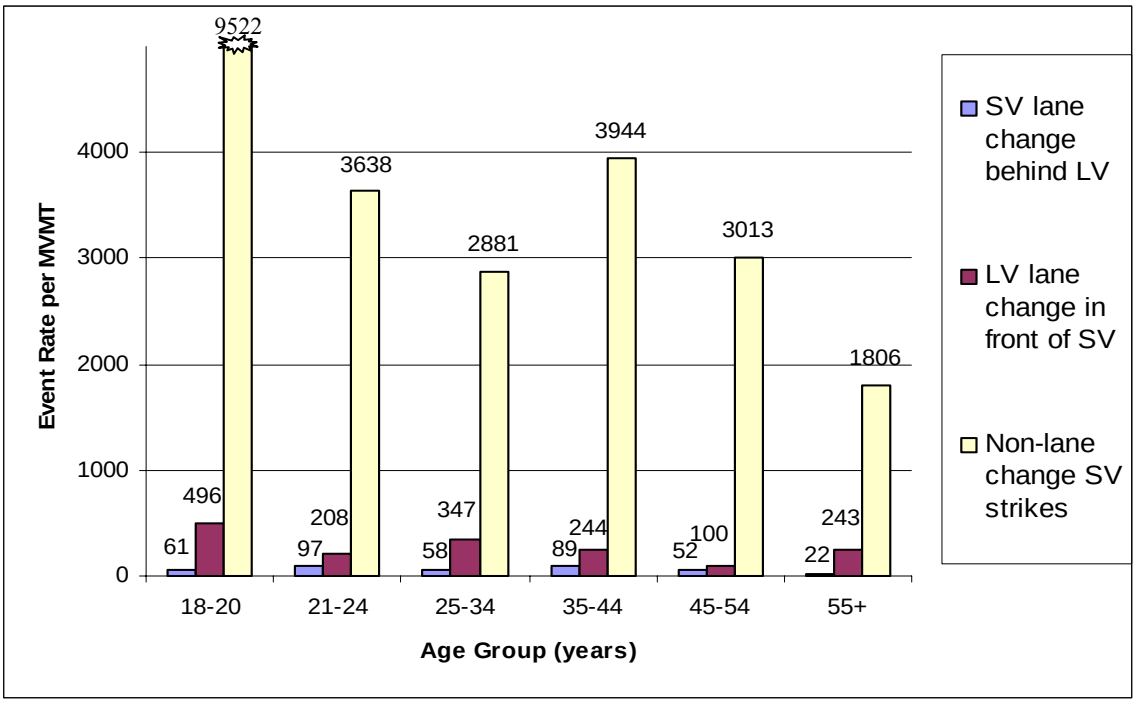


Figure 10.1. Rate per MVMT for SV striking by age and lane change maneuver for incidents.

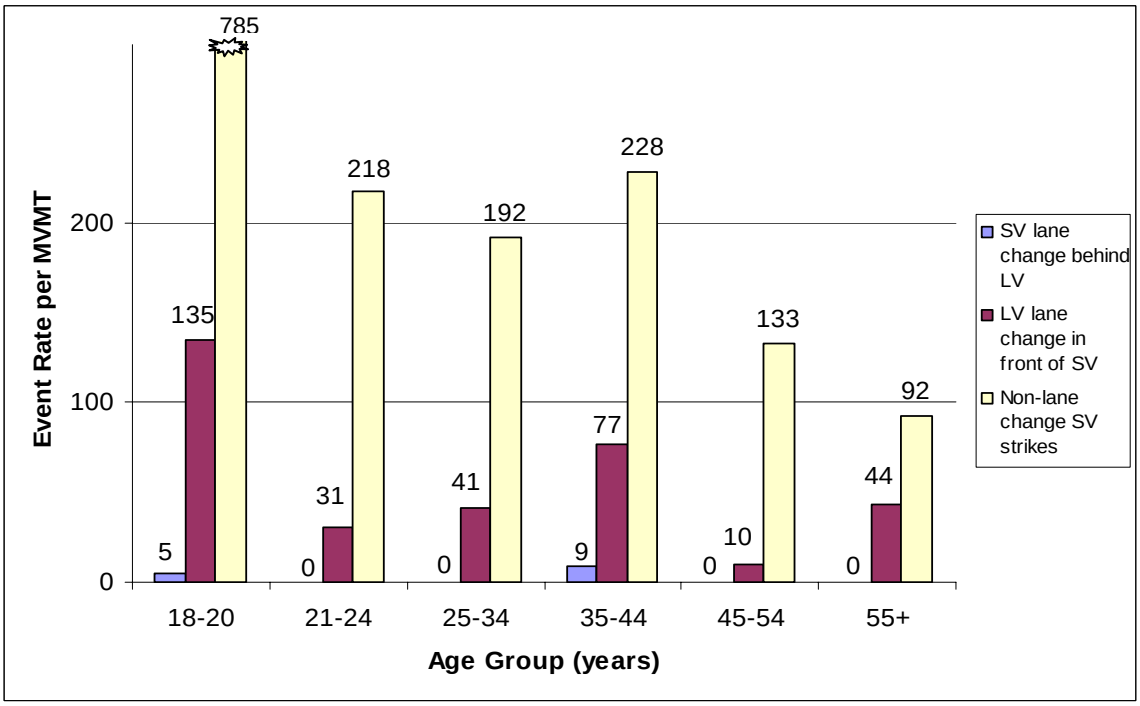


Figure 10.2. Rate per MVMT for SV striking by age and lane change maneuver for near-crashes.

Figure 10.3 presents the rate data for SV striking incidents by both age and gender. In order to simplify the presentation of results and the discussion, age categories were combined to make a total of three age groupings: 18-24; 25-44 and; 45+. The only clear pattern that emerges from this figure is for the 45+ age grouping. The male drivers in this age group had a rate that was more than twice as high as that for female drivers for both lane change scenarios (*SV lane change behind LV* and *LV lane change in front of SV*). The gender comparisons were fairly even across the other age groupings. Almost all of the near-crashes were of the *LV lane change in front of SV* type (Figure 10.4). Males had a noticeably higher near-crash rate than females for this scenario in the 18-to-24 and 45+ age groups, while females had a higher rate in the 25- to-44 age group.

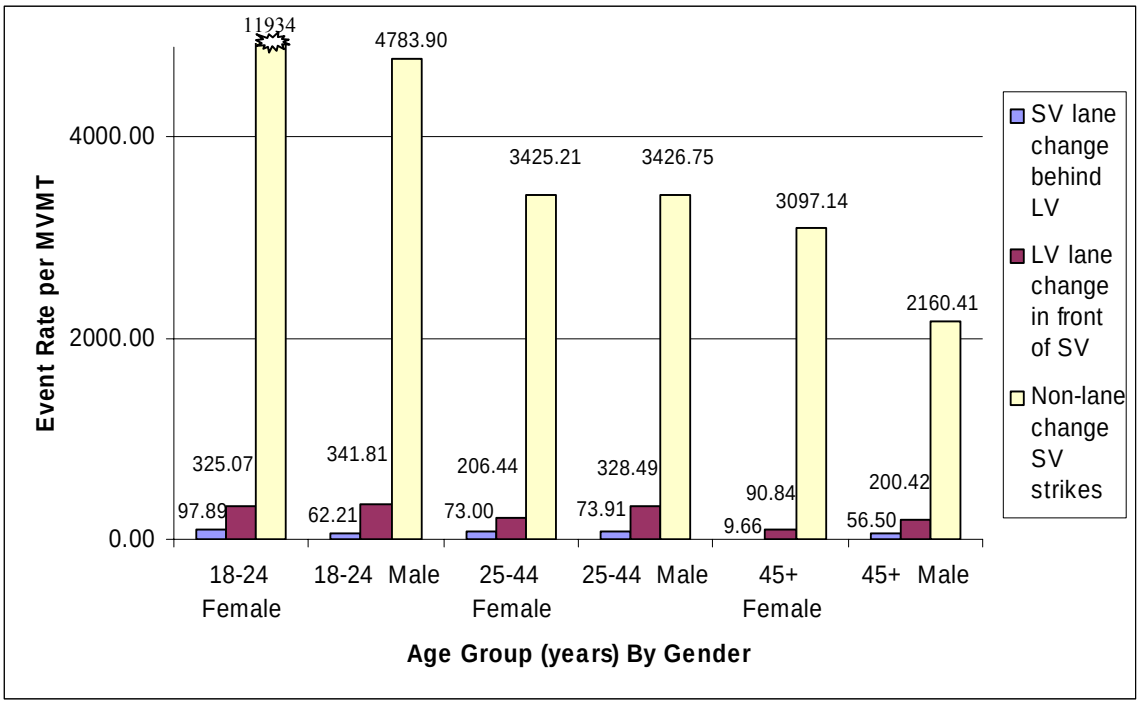


Figure 10.3. Rate per MVMT for SV striking by age and by gender by lane change maneuver for incidents.

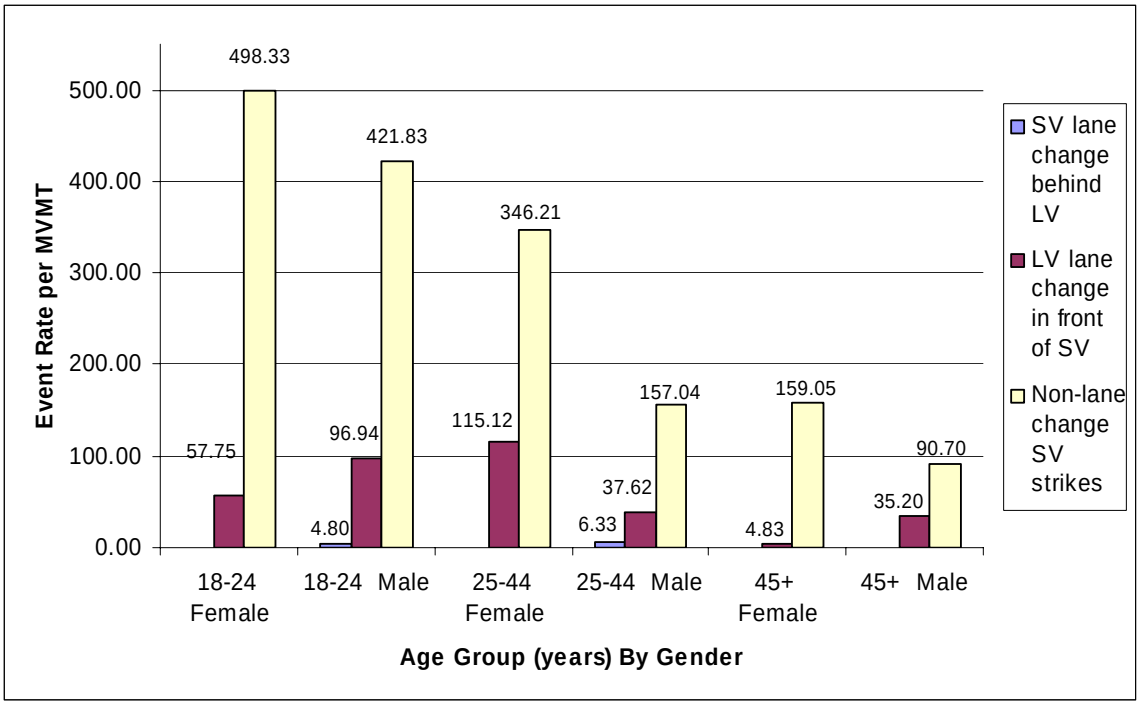


Figure 10.4. Rate per MVMT for SV striking by age and gender by lane change maneuver for near-crashes.

SV Struck Data

Incident rates by age group for SV struck incidents are shown in Figure 10.5 by age group. Nearly all of the lane change incidents were of the *SV lane change in front of FV* type. Three age categories (18-20, 25-34, and 35-44) had incident rates that were 1.5 to 3.5 times as high as the other three age groups. Except for the 21- to 24-year-olds, there was a downward trend with increasing age. For near-crashes, shown in Figure 10.6, the three younger age groups had rates 1.5 to 4 times as high as the three older age groups.

Age and gender were considered next for the SV struck events. Again, only the *SV lane change in front of FV* scenario will be considered, as nearly all the incidents and near-crashes were of this type. As seen in Figure 10.7, females had nearly twice the rate as males for each of the three age groups. A similar pattern was observed for near-crashes, except that the 18- to 24-year-old males and females had virtually identical rates and the remaining differences for the other age groups were by at least a 4 to 1 margin (females higher than males).

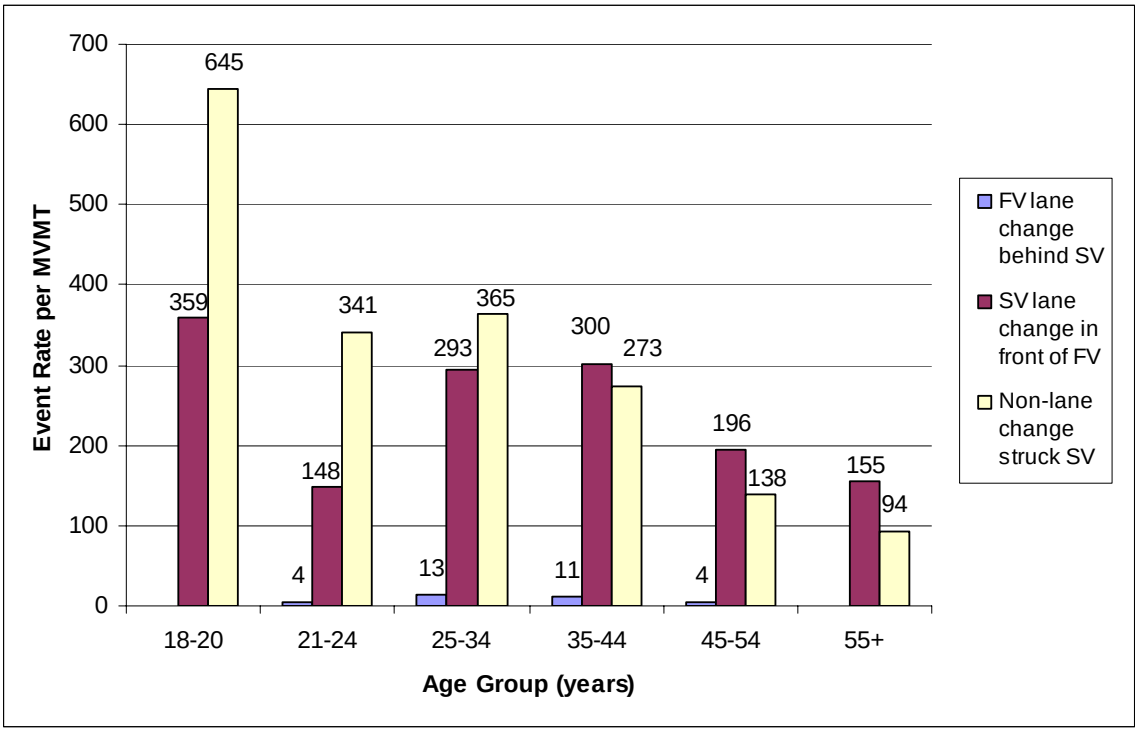


Figure 10.5. Rate per MVMT for struck SV by age and lane change maneuver for incidents.

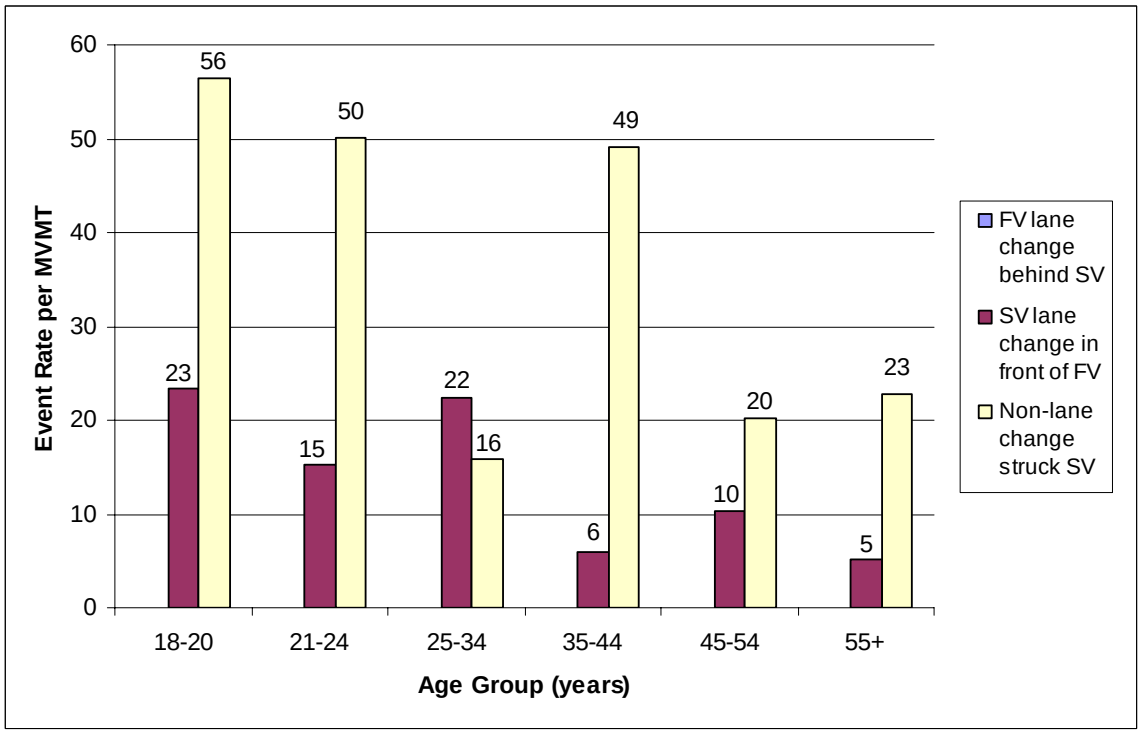


Figure 10.6. Rate per MVMT for struck SV by age and lane change maneuver for near-crashes.

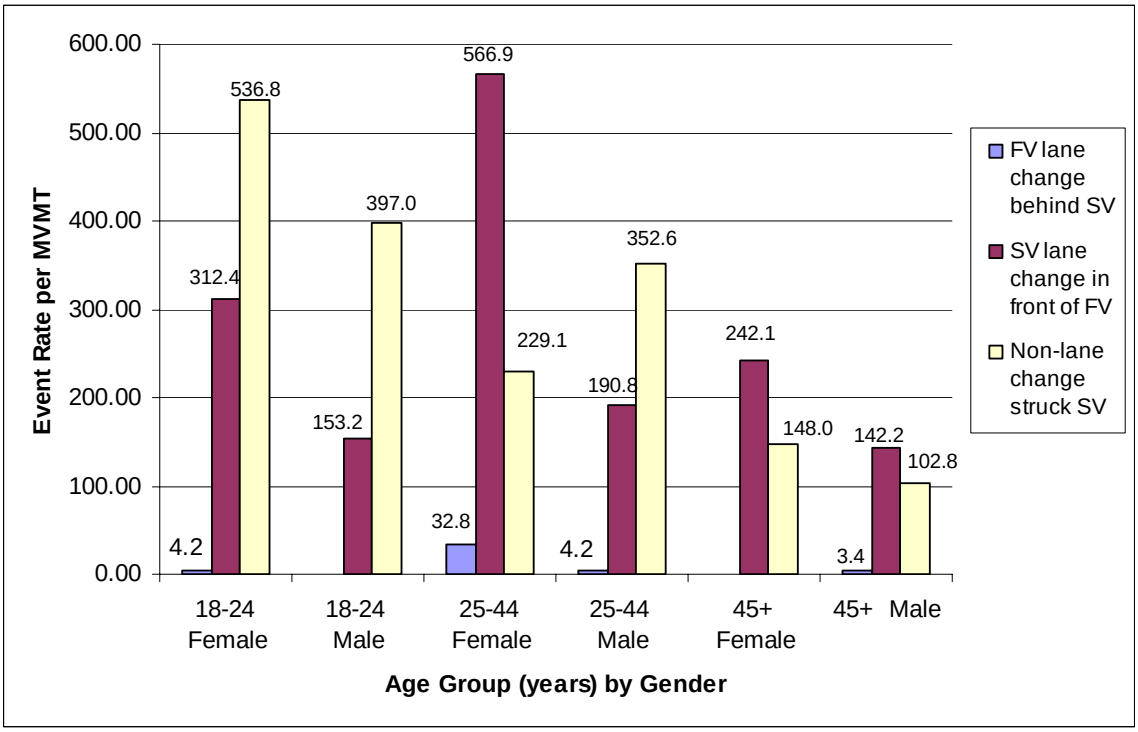


Figure 10.7. Rate per MVMT for struck SV by age and gender by lane change maneuver for incidents.

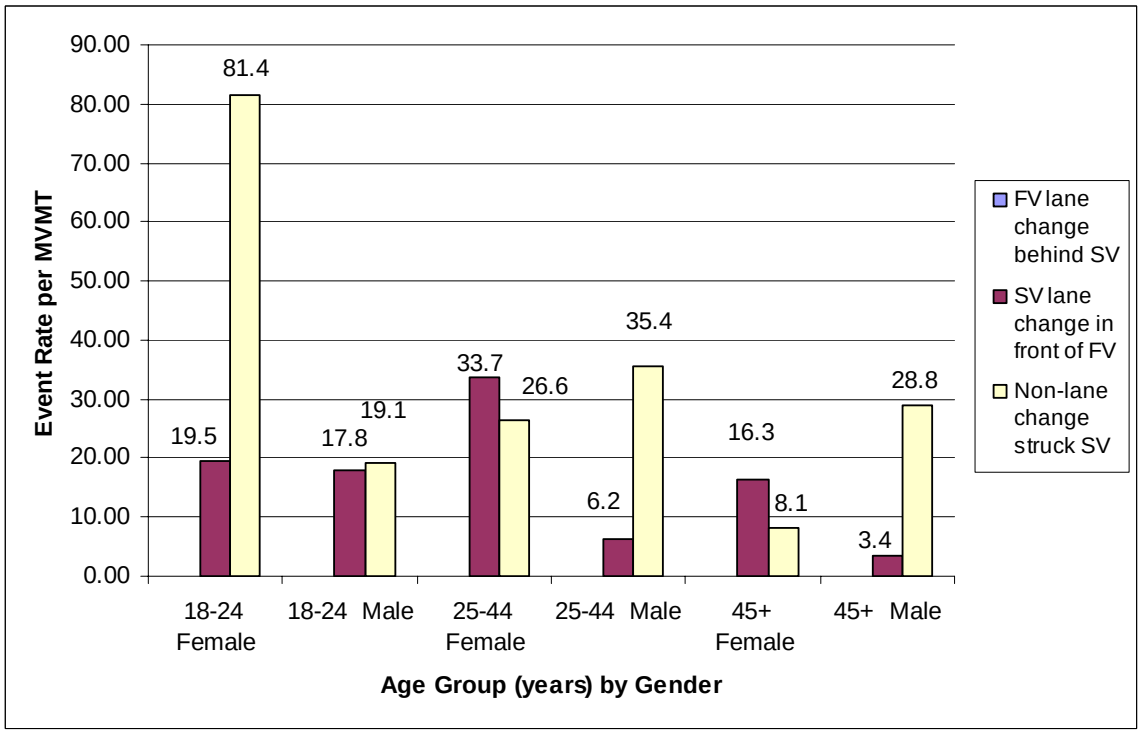


Figure 10.8. Rate per MVMT for struck SV by age and gender by lane change maneuver for near-crashes.

Question 2. What are the primary contributing factors associated with these RE conflicts associated with the lane-change-related initial conditions?

Question 2 examined the contributing factors for lane change conflicts. As for previous analyses in this chapter, rate per MVMT was examined. Two types of contributing factors were considered.

Driver factors considered in these analyses included:

- Driver Physical/Mental Impairment
- Driver 1 Distracted By
- Willful Behavior
- Driver Proficiency
- Driver 1 Vision Obscured by

Environmental and roadway factors considered included:

- Roadway Infrastructure
- Relation to junction
- Traffic control (counted if traffic control present)
- Roadway Alignment (counted if anything other than straight)
- Weather (counted if anything but sunny)
- Surface Condition (counted if anything but dry)
- Light (counted if anything other than day)
- Traffic density (counted if anything other than free flow)

As before, only incidents and near-crashes will be presented and discussed, since there were no lane-change-related crashes. Rate data per MVMT will be used in all cases. Environmental and roadway factors will be discussed first, followed by driver factors.

Environmental and Roadway Factors

When the incident data were analyzed (Table 10.3), *traffic density* was the factor with the highest rate for all four lane-change-related scenarios, by a factor of at least four, in all but one case. Going back to the frequency data, 90 percent of all lane change-related incidents were coded with *traffic density* as a contributing factor. *Light*, *traffic control*, and *relation to junction* were second, third, or fourth most important for all four scenarios. For the non-lane-related incidents, the highest rates were observed for *traffic density*, *relation to junction*, *traffic control*, and then *light*, quite a different pattern than was seen for the lane change-related incidents.

When the near-crash data were considered, as shown in Table 10.4, the lane-change-related scenarios followed similar patterns as they did for incidents. For near-crashes, the non-lane-related rate pattern was more closely aligned to the lane-change-related near-crash data than to the non-lane-related incident data.

Table 10.3. Rate of incidents per MVMT for SV struck and SV striking environmental and roadway factors.

Environmental and Roadway Factors	SV Striking			SV Struck		
	SV lane change behind LV	LV lane change in front of SV	Non-lane change SV strikes	FV lane change behind SV	SV lane change in front of FV	Non-lane change struck SV
Relation to Junction	11.18	79.06	1,702.26	3.82	47.44	127.39
Traffic Control	10.21	49.91	1,333.21	2.08	35.12	110.32
Roadway Alignment	9.53	24.02	362.00	0.51	9.57	27.41
Weather	11.22	34.99	476.10	0.51	31.35	45.50
Surface Condition	2.13	33.57	330.74	0.51	21.10	17.23
Light	19.85	75.02	1,020.39	1.04	73.64	62.28
Traffic Density	61.90	248.72	3,742.11	5.81	214.81	270.47

Table 10.4. Rate of near-crashes per MVMT for SV struck and SV striking environmental and roadway factors.

Environmental and Roadway Factors	SV Striking			SV Struck		
	SV lane change behind LV	LV lane change in front of SV	Non-lane change SV strikes	FV lane change behind SV	SV lane change in front of FV	Non-lane change struck SV
Relation to Junction	1.25	15.45	82.98	0.00	2.68	11.99
Traffic Control	1.25	6.96	61.45	0.00	1.08	9.70
Roadway Alignment	0.50	12.54	37.36	0.00	0.72	5.40
Weather	1.25	11.75	54.87	0.00	4.88	8.71
Surface Condition	0.75	2.95	42.06	0.00	2.54	4.82
Light	0.50	13.28	102.19	0.00	7.37	6.25
Traffic Density	2.38	38.73	215.91	0.00	10.05	29.30

Driver Factors

The driver factors referred to the driver of the SV, whether the SV was in the role of the striking or stuck vehicle. Driver factors such as willful behavior and driver proficiency were coded for lane change-related incidents such as when the driver braked abruptly for no apparent reason. Another example of SV behavior putting the SV at greater risk of becoming the struck vehicle was leaving an insufficient gap when changing lanes. This sort of behavior could violate the expectations of other drivers and create conflicts with the following vehicle. The driver factors for all lane change-related incidents are shown in Table 10.5. Driver proficiency showed up as a prominent factor for the *SV struck* scenarios (higher than the next highest driver factor rate by 4 to 1). For the *SV striking* scenarios, the incidents rates for driver factors were fairly even within each scenario. Here, driver proficiency was the second highest for the *SV lane change behind LV* scenario (the highest rated factor was willful behavior). When the near-crash rates in Table 10.6

were examined, driver proficiency and driver distraction were also the factors with the highest rates.

Table 10.5. Rate of incidents per MVMT for SV struck and SV striking driver factors.

Driver Factors	SV Striking			SV Struck		
	SV lane change behind LV	LV lane change in front of SV	Non-lane change SV strikes	FV lane change behind SV	SV lane change in front of FV	Non-lane change struck SV
Driver Physical/Mental Impairment	1.79	22.89	435.58	0.00	4.51	28.03
SV Driver Distracted By	10.55	43.82	1,033.03	0.00	21.26	40.69
Willful Behavior	31.20	24.33	443.47	0.00	40.05	20.35
Driver Proficiency	27.89	66.76	2,496.71	0.00	162.13	115.87
SV Driver Vision Obscured By	4.31	14.12	365.39	0.00	12.37	21.68

Table 10.6. Rate of near-crashes per MVMT for SV struck and SV striking driver factors.

Driver Factors	SV Striking			SV Struck		
	SV lane change behind LV	LV lane change in front of SV	Non-lane change SV strikes	FV lane change behind SV	SV lane change in front of FV	Non-lane change struck SV
Driver Physical/Mental Impairment	1.12	5.50	40.54	0.00	0.00	4.84
SV Driver Distracted By	0.50	4.48	123.56	0.00	0.00	11.44
Willful Behavior	2.38	7.73	40.85	0.00	0.00	4.34
Driver Proficiency	0.75	4.47	155.24	0.00	0.00	15.00
SV Driver Vision Obscured By	1.01	1.06	40.70	0.00	0.00	2.44

Question 3. What are the frequencies of lane change conflicts for side-sensor--equipped vehicles and what were the dynamic conditions (vehicle traveling in blind spot, approaching vehicle in adjacent lane) associated with these conditions?

As stated earlier in the chapter, 20 of the leased vehicles were retrofitted with side sensors for the last six months of the study. For the side-sensor-equipped vehicles, data were triggered as follows:

- Side Cutoff: For this trigger, vehicle speed had to be greater than 8.9 m/s (20 mph) and a lane change occurred in front of another car located within 15.2 m (50 ft) of the subject vehicle. Events validated for this trigger were successful lane changes in which a vehicle traveling in the adjacent lane at the beginning of the maneuver was cut off by the SV.

- Turn signal: This trigger occurred when an object was detected by the side radar within +/- 1 second of any instance in which the turn signal light was active. Vehicle speed also had to be higher than 6.7 m/s (15 mph) for the trigger to occur. Events validated for this trigger were lane change aborts (SV driver obviously wanted to change lanes, but was prevented from doing so by the presence of a vehicle in the adjacent blind spot).
- Side Blind Spot: This trigger occurred only for vehicle speeds greater than 8.9 m/s (20 mph) when a lane abort maneuver (as detected by the lane tracker) occurred while an object was detected by the side radar. Events validated for this trigger were lane change aborts as defined above.
- Side Yaw: The trigger criterion for yaw rate was any set of values that went from neutral (i.e., ~0) yaw rate to +2 degrees/sec, oscillated back to -2 degrees/sec (or vice versa: -2 to +2), and then returned to neutral within a 3-second time window. A minimum speed of 6.7 m/s (15 mph) was required for the trigger to activate. In addition, the side radar had to be detecting an object for the trigger to occur. Events validated for this trigger were lane change aborts as defined above.

These triggers helped identify a total of 19 lane change abort events (vehicle traveling in blind spot) and 261 lane change cutoff events (vehicle traveling in the adjacent lane at the beginning of the maneuver was cut off by the SV) as shown in Table 10.7. The ratio of successful cutoff lane changes to lane change aborts was 13.8 to 1. With 20 vehicles and 6 months of data collection, there were 120 months of side radar data collected, resulting in 2.3 lane change events identified per vehicle-month (an average of 14 conflicts identified for each leased vehicle over the six-month time frame). Four events are not presented in Table 10.7. In two cases, there were threats from both the left and right (a quick swerve around a stopped vehicle with oncoming traffic present). I n the other two cases, the threat was from a guardrail rather than another vehicle; both occurred when the SV was traveling too fast on an exit ramp curve and went over the lane line towards the guardrail such that it was perceived by the side radar as a threat.

Table 10.7. Lane change conflict frequencies by direction and type for side sensor-equipped vehicles.

Type of Lane Change Conflict	Left	Right
Lane change aborts	9	10
Cutoff vehicle in adjacent lane	166	95

Figure 10.9 shows a scatterplot of range and range-rate for cutoff lane change conflicts. All cutoff events but one occurred with an initial range of less than 15.5 m (51 feet). Events of the left half of the graph (negative range rate) indicate cases in which the SV was initially traveling slower than the cutoff vehicle. Note that distances are generally longer for these cases. That is, if the SV had a negative range rate in relation to the cutoff vehicle, it would be considered a valid event at a greater distance than would be true for the same event with a positive range rate. For some events with a positive range rate (SV pulling away from the cutoff vehicle), there was also a forward threat that the SV was closing on; some of the longer range events on the right side of the graph fell into this category. That is, even though the SV was pulling away from the cutoff

vehicle and was a decent distance from the cutoff vehicle, there was another threat (e.g., lead vehicle) that made this a valid event.

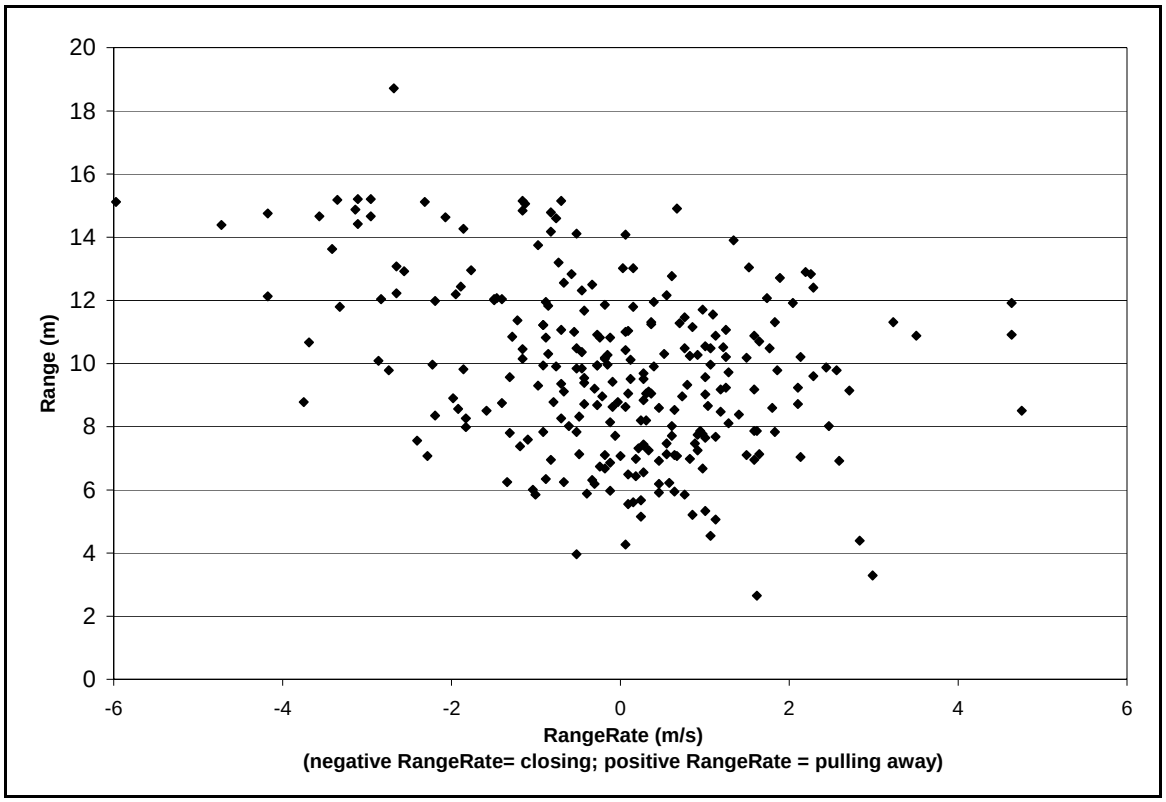


Figure 10.9. Range and range-rate values for cutoff lane change conflicts.

DISCUSSION

A primary overall goal for the 100-Car Study was to determine the causes and contributing factors associated with RE crashes. An important aspect of fully understanding the rear-end crash problem is understanding the pre-event maneuvers and precipitating factors that, in conjunction with other contributing factors, lead to RE crashes. The purpose of the analyses for Chapter 10, *Goal* 6 was to understand the degree to which lane change events, such as cut-ins, lead to rear-end conflicts. This has important implications for the design of future forward collision warning systems, since a cut-in vehicle may not provide a radar signature until very late in a conflict scenario. To begin to understand this issue, the RE conflict data were analyzed for both lead-vehicle (i.e., subject vehicle as following vehicle) and following-vehicle (i.e., subject vehicle as lead vehicle) scenarios. Frequency distributions were generated to identify the rate that these types of scenarios occurred per MVMT, the initial kinematic conditions that occurred for each, and the contributing factors that played a role for each type of scenario.

No crashes occurred when there was a lane change as a precipitating factor in front of the subject vehicle or when there was a lane change behind the subject vehicle. There were, however, 64 near-crashes and 324 incidents that occurred when there was a cut-in to the lane in front of the

subject vehicle as compared to only 4 near-crashes and 77 incidents when the SV changed behind a lead vehicle. As will be described in Chapter 11, *Goal 7*, the subject vehicle drivers were judged to more often be impaired (30 incidents more), distracted (44 incidents more) and make proficiency-related errors (e.g., inappropriate reaction; 55 more) in the cases when they were cut-off. This seems to support what has been found throughout this report. At least two elements are required for a conflict to occur and most often requires a precipitating maneuver plus another contributing factor (often driver state-related) for an event to occur. In this case, there were fewer events when the subject vehicle was the cut-in vehicle because the drivers were presumably alert and attentive when they were actively performing the lane change maneuver.

The lane-change-related SV striking events were analyzed according to age grouping. For lane change incidents, the 18- to 20-year-olds had the highest rate for *SV lane change behind LV*, while the other age categories had fairly equal rates for this scenario. For the *LV lane change in front of SV* scenario, the 21- to 24-year-olds had the highest rate, although the rates were fairly even across age groupings. For near-crashes, 18- to 20-year-olds again had the highest rate for the *SV lane change behind LV* scenario by a factor of 2 to 1 over the next highest age group (35-44-year-olds). Rate data for SV striking events by both age and gender were considered next. The only clear pattern that emerged from this figure was for the 45+ age grouping. The male drivers in this age group had a rate that was more than twice as high as that for female drivers for both lane change scenarios (*SV lane change behind LV* and *LV lane change in front of SV*). The gender comparisons were fairly even across the other age groupings. Almost all of the near-crashes were of the *LV lane change in front of SV* type (Figure 10.4). Males had a noticeably higher near-crash rate than females for this scenario in the 18- to 24 and 45+ age groups, while females had a higher rate in the 25- to 44 age group.

For lane-change-related SV struck events, three age categories (18-20, 25-34, and 35-44) had incident rates that were 1.5 to 3.5 times as high as the other three age groups. Except for the 21- to 24-year-olds, there was also a downward trend in events with increasing age. For near-crashes, the three younger age groups had rates 1.5 to 4 times as high as the three older age groups. When both age and gender were considered, females had at least twice the rate as males for each of the three age groups. A similar pattern was observed for near-crashes, except that the 18- to 24-year-old males and females had virtually identical rates and the remaining differences for the other age groups were by at least a 4 to 1 margin (females higher than males).

Roadway and infrastructure factors were considered next. *Traffic density* was the factor with the highest incident rate for all 4 lane-change-related scenarios, by a factor of at least four in all but one case. Going back to the frequency data, 90 percent of all lane-change-related incidents were coded with *traffic density* as a contributing factor. *Light*, *traffic control*, and *relation to junction* were second, third, or fourth most important for all 4 scenarios. For the non-lane-related incidents, the highest rates were observed for *traffic density*, *relation to junction*, *traffic control*, and then *light*, quite a different pattern than was seen for the lane-change-related incidents. When the near-crash data were examined, the lane-change-related scenarios followed similar patterns as for incidents. For near-crashes, the non-lane-related rate pattern was more closely aligned to the lane-change-related near-crash data than to the non-lane-related incident data.

Driver contributing factors were also examined. Driver proficiency showed up as a prominent factor for the *SV struck* incident scenarios (higher than the next highest driver factor rate by 4 to 1). For the *SV striking* scenarios, the incident rates for driver factors were fairly even within each scenario, and driver proficiency was not even the top factor for the *SV lane change behind LV* scenario (the highest rated factor was willful behavior). When the near-crash rates were examined, driver proficiency and driver distraction also had the highest rates.

Twenty of the leased vehicles were retrofitted with side sensors for the last six months of the study. Four triggers based on the side radar and turn signal, the lane tracker, and the yaw sensor helped identify a total of 280 lane change abort (vehicle traveling in blind spot) and lane change cutoff (vehicle traveling in the adjacent lane at the beginning of the maneuver was cut off by the SV) events. The ratio of successful cutoff lane changes to lane change aborts was 13.8 to 1. All cutoff events but one occurred with an initial range of less than 15.5 m (51 feet). Events with an initial negative range rate indicated cases in which the SV was initially traveling slower than the cutoff vehicle; ranges were generally longer for these cases. That is, if the SV had a negative range rate in relation to the cutoff vehicle, it would be considered a valid event at a greater distance than would be true for the same event with a positive range rate. For some events with a positive range rate (SV pulling away from the cutoff vehicle), there was also a forward threat that the SV was closing on. That is, even though the SV was pulling away from the cutoff vehicle and was a decent distance from the cutoff vehicle, there was another threat present that made this a valid event.

This page intentionally left blank.

CHAPTER 11: GOAL 7, DETERMINE THE DISTRIBUTION OF INATTENTION FOR EACH REAR-END LEAD-VEHICLE SCENARIO

DATA ANALYSIS OVERVIEW

The purpose of this goal is to analyze the impact of driving inattention on rear-end striking crashes, near-crashes, and incidents. For a more general treatment of driver inattention, refer to Chapter 7, *Goal 3*. For both this chapter and Chapter 7, *Goal 3*, driver inattention is operationally defined somewhat differently for crashes, near-crashes, and incidents. For *crashes* and *near-crashes,* inattention is considered to be present when drowsiness, driving-related inattention to forward roadway, secondary task performance, or nonspecific eyeglance away from the forward roadway was identified as a contributing factor either during the initial data reduction or during the eyeglance data reduction. For *incidents*, inattention is present when drowsiness, driving-related inattention to forward roadway, or secondary task performance was identified as a contributing factor during initial data reduction. These separate definitions were necessary because project resources did not allow eyeglance data reduction for the large number of incidents analyzed for this study.

Recall from Chapter 7, *Goal 3* that reductionists recorded drowsiness as a contributing factor for drivers who exhibited drowsy behaviors, including slow eyelid closures (Wierwille and Ellsworth, 1994). These events represent moderate to severe drowsiness, and are grouped under the general heading of driver inattention since they are unable to devote a "safe" level of attention to driving due to eyelid closures or the general effects of drowsiness. Driving-related inattention to the forward roadway included those events when the driver was performing a driving-relevant task (i.e., checking the speedometer, checking rear-view mirrors, blind spots, or looking 90 deg to check for cross traffic), but was not looking at the forward roadway. Secondary tasks were recorded by the data reductionists when drivers were engaging in specified behaviors within 3 seconds of the onset of the precipitating factor for each event (see Appendix D for a list of these behaviors). Finally, nonspecific eyeglance away from the forward roadway was noted for instances when the driver had at least one glance from the roadway just prior to, or during the onset of, the precipitating factor. A nonspecific eyeglance included a glance at a location that was toward a non-discernable object either inside or outside the vehicle. Events for this category of inattention were derived from the eyeglance analysis and are thus available only for the crashes and near-crashes. In addition, more than one type of inattention was listed for some events; therefore, combinations of inattention types will be listed for some of the figures in this chapter.

Five different lead-vehicle kinematic scenarios identified during data reduction were found to describe a majority of the lead-vehicle conflicts. These scenarios are as follows:

- Lead vehicle stopped greater than 2 seconds.
- Lead vehicle stopped less than or equal to 2 seconds.
- Lead vehicle decelerating.
- Lead vehicle moving at a slower, constant speed.
- Lead vehicle accelerating, but traveling at a slower rate.

Data Included in the Analyses

For the following analyses, all 241 drivers were used in the frequency counts, since the focus of this research goal was to evaluate all inattention-related lead-vehicle events by the 5 kinematic scenarios. With the data reduced to this level of detail, age and gender analyses resulted in many empty cells. Therefore, all drivers (both primary and secondary) were used, and no reduction by age, gender, or vehicle miles traveled was conducted for this research objective. Please note that only the lead-vehicle conflicts where the precipitating event was one of the 5 kinematic scenarios listed above were used in these analyses.

Question 1. What is the frequency of each RE scenario for which inattention was a factor as compared to those instances for which inattention was not a factor?

The purpose of this question is to examine the effect of inattention on lead-vehicle conflicts and compare the occurrence of inattention to those conflicts when the driver remained attentive. Event severity (crash, near-crash, or incident) is compared for each lead-vehicle scenario. Note that the information presented below counts the number of events when at least one type of inattention occurred. If multiple inattention tasks occurred in the same event, the combinations of inattention are shown. Also note that evaluating the data at this level resulted in many cells with missing values, therefore, only frequencies and percentages are presented.

Figure 11.1 shows the total number of crashes, near-crashes, and incidents associated with at least one of the four types of inattention versus those crashes, near-crashes and incidents where the driver remained attentive to the forward roadway. For crashes, those marked as inattentive outnumbered those with an attentive driver by 13 to 1. For near-crashes, the ratio is reduced to approximately 2 to 1. For incidents, this trend is reversed and attentive incidents outnumbered inattentive incidents by 2 to 1.

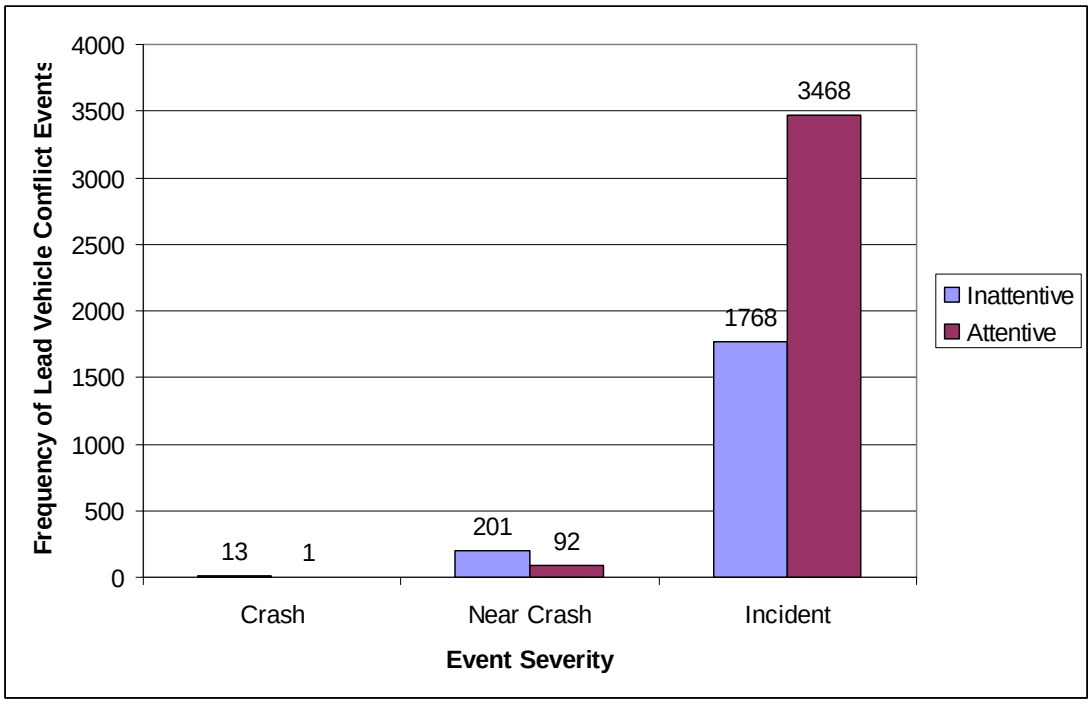


Figure 11.1. Frequency of crashes, near-crashes, and incidents where inattention was a factor.

Since there were many more incidents than crashes, Figure 11.2 shows these data in terms of percentage for each severity level in order to bring out the interaction between inattention and event severity (i.e., percentage of LV crashes marked as inattentive *plus* the percentage of LV crashes marked as inattention *is equal to* the total number of LV crashes). Note that approximately one-third of the incidents have inattention listed as a contributing factor whereas 93 percent of the crashes are inattention-related. The effect is nearly perfectly linear, and seems to indicate a strong correlation between inattention and increased severity for lead-vehicle RE conflicts.

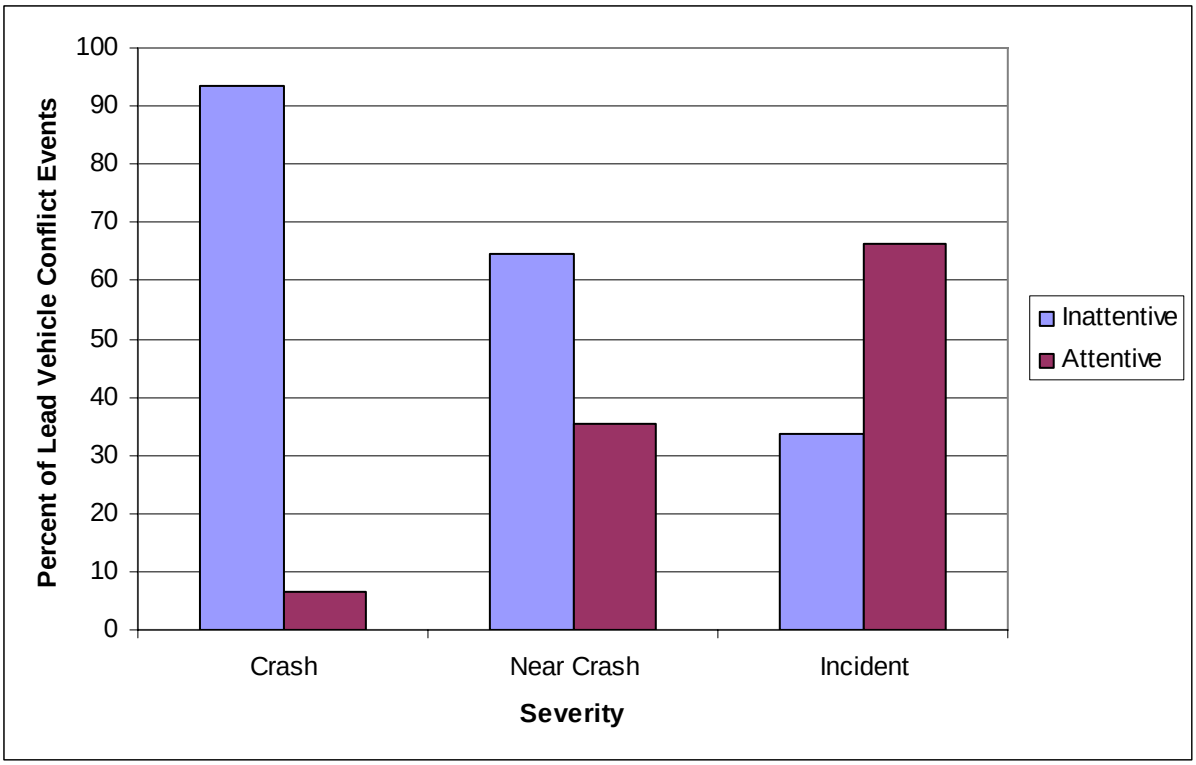


Figure 11.2. Percentage of lead-vehicle crashes, near-crashes, and incidents for which inattention was listed as a contributing factor.

Figure 11.3 shows the total number of crashes and near-crashes associated with the four inattention categories as well as the combinations of inattention categories. Two items are worth noting. First, combinations of different types of inattention are generally low in number, except for *secondary task + nonspecific eyeglance away from the forward roadway*. Second, eyeglances away from the forward roadway contributed to four of the 5 most frequent types of inattention for both crashes and near-crashes (*driving-related inattention to the forward roadway*, *secondary task + nonspecific eyeglance*, *drowsiness,* and *nonspecific eyeglance away from the forward roadway*).

Figure 11.4 shows the percentage of the inattention categories for crashes and near-crashes. An important finding of this research is that inopportune eyeglances (those that are close in time to the precipitating factor) are a primary contributing factor for crashes and near-crashes.

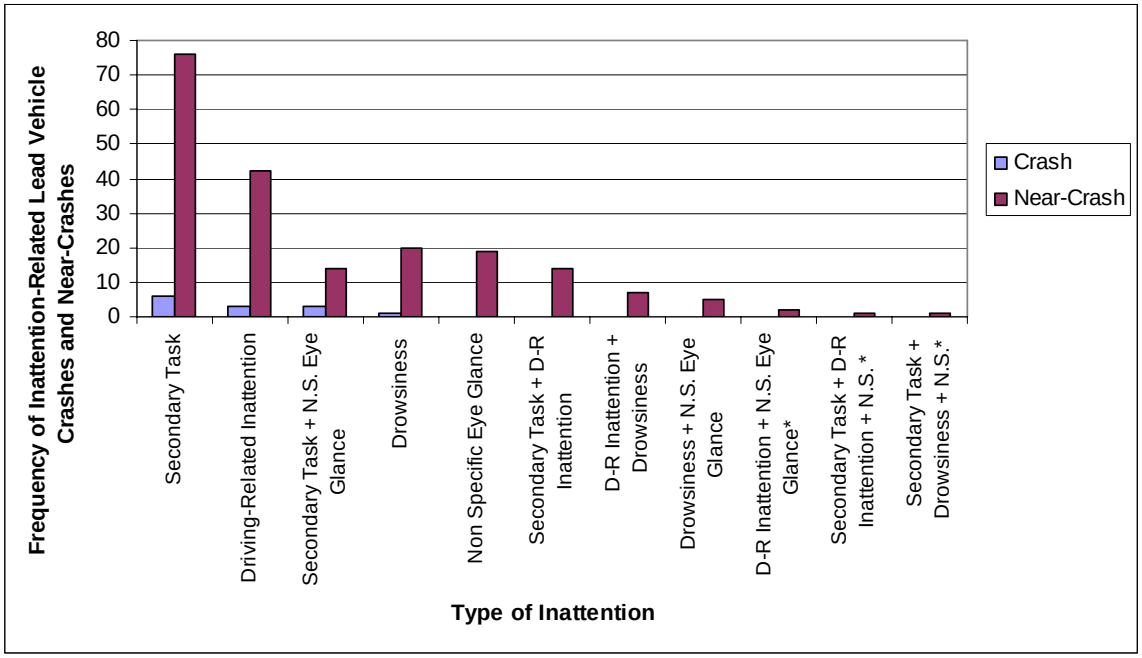


Figure 11.3. The frequency of inattention categories for crashes and near-crashes.

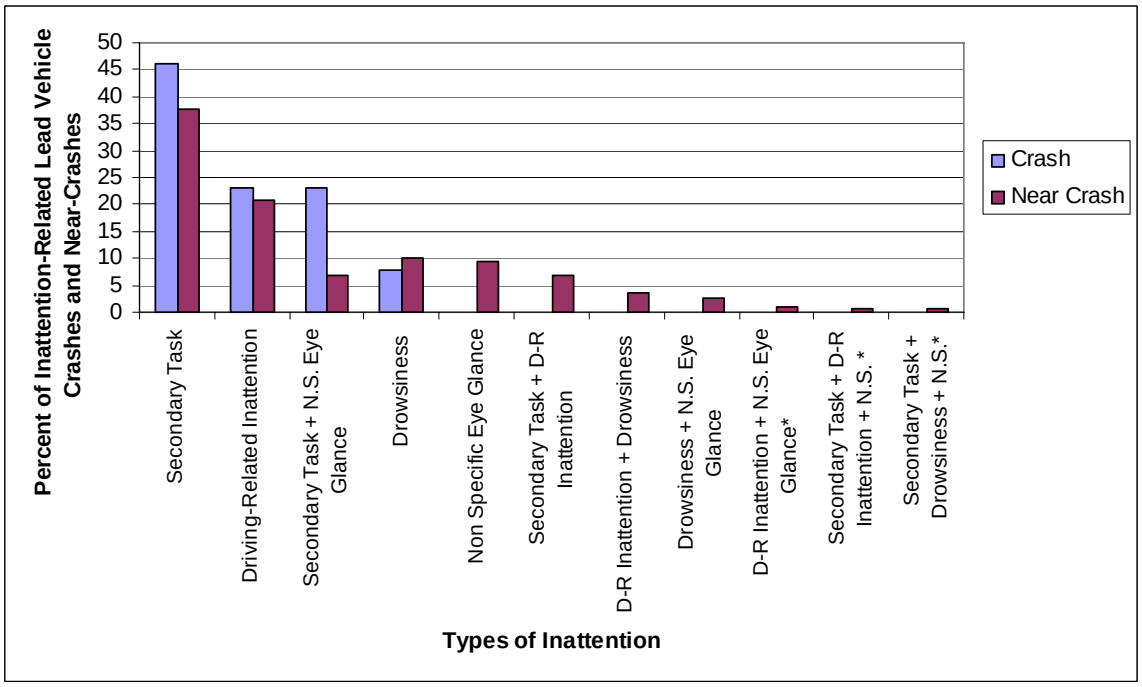


Figure 11.4. The percentage of inattention categories for crashes and near-crashes.

Figure 11.5 shows the number of incidents associated with each type of inattention. Once again, the combinations of inattention types comprised a small number of the total number of incidents. The most frequent type of inattention was *secondary task engagement*, followed by *drowsiness* and *driving-related inattention to the forward roadway*. Future reports will incorporate the

eyeglance analyses (*nonspecific eyeglance away from the forward roadway*) for incidents, thus providing greater comparability to the crash and near-crash numbers.

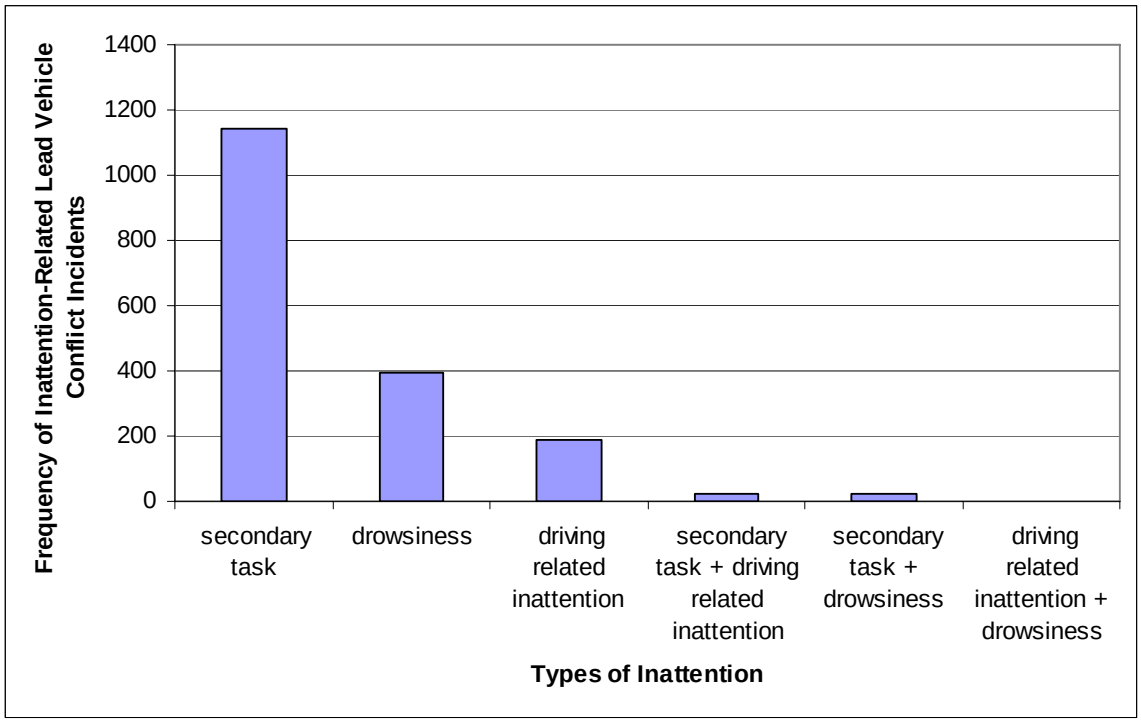


Figure 11.5. Frequency of inattention-related lead-vehicle conflict incidents by type of inattention or combination of types of inattention.

Lead-Vehicle Scenario Analysis

The following analysis discusses the effect of inattention on all lead-vehicle conflicts and demonstrates whether inattention may have affected the severity of the events. For these figures, events with at least one type of inattention are included in the frequency count. Therefore, regardless of how many types of inattention may have contributed to these events, each event is only counted once.

Figure 11.6 shows lead-vehicle crashes by lead-vehicle scenario type. Crashes only occurred for the two stopped LV scenarios (LV stopped > 2 s and LV stopped $\leq$ 2 s). Of the 14 crashes shown, inattention was marked as a contributing factor for 13 of lead-vehicle crashes (93%).

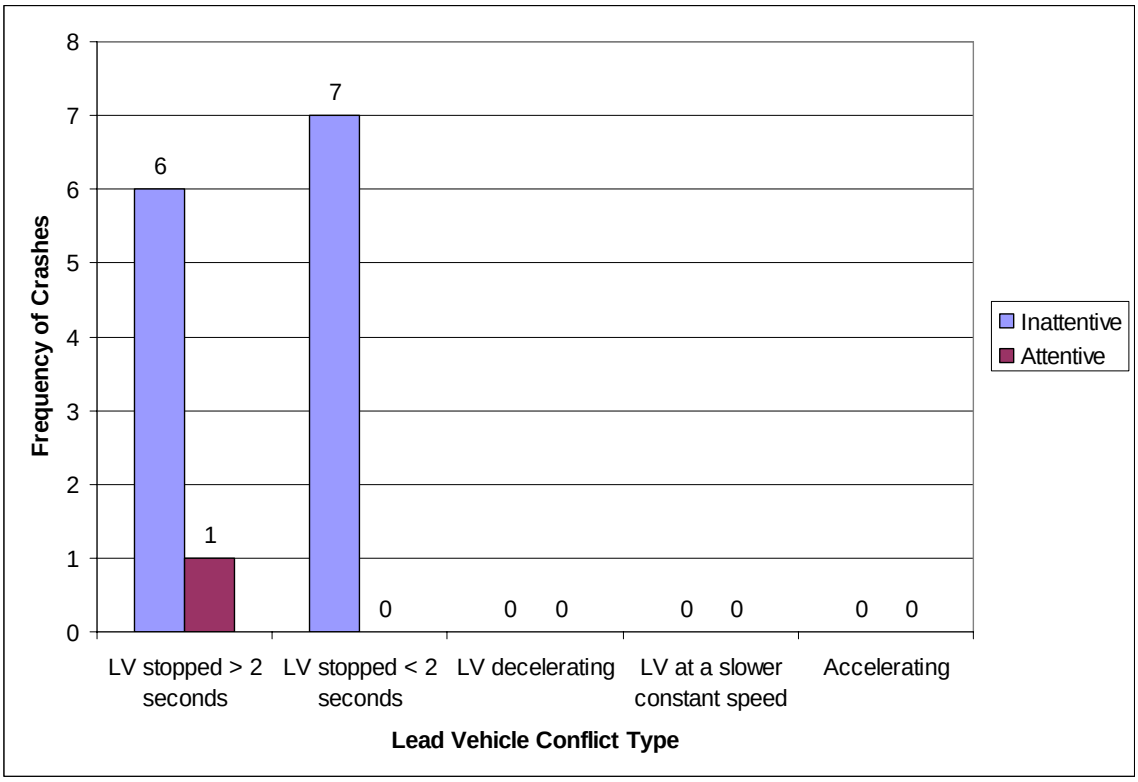


Figure 11.6. Frequency of crashes by driver attention and lead-vehicle kinematic scenario.

Figure 11.7 presents the lead-vehicle near-crashes by lead-vehicle scenario type and level of attention. For near-crashes, inattention was a factor almost 70 percent of the time, although this varied to some extent by LV scenario. For example, inattention was a factor in 59 percent of *LV stopped > 2s*, 73 percent of *LV stopped ≤ 2s*, and 68 percent of *LV decelerating* near-crashes. These results indicate that the role of inattention in near-crashes may vary according to the LV scenario.

Figure 11.8 presents the lead-vehicle incidents by lead-vehicle scenario type and level of attention. Overall, inattention was a factor more than 34 percent of the time for incidents. This percentage stayed fairly constant (at about 32 to 36%) for each LV scenario examined.

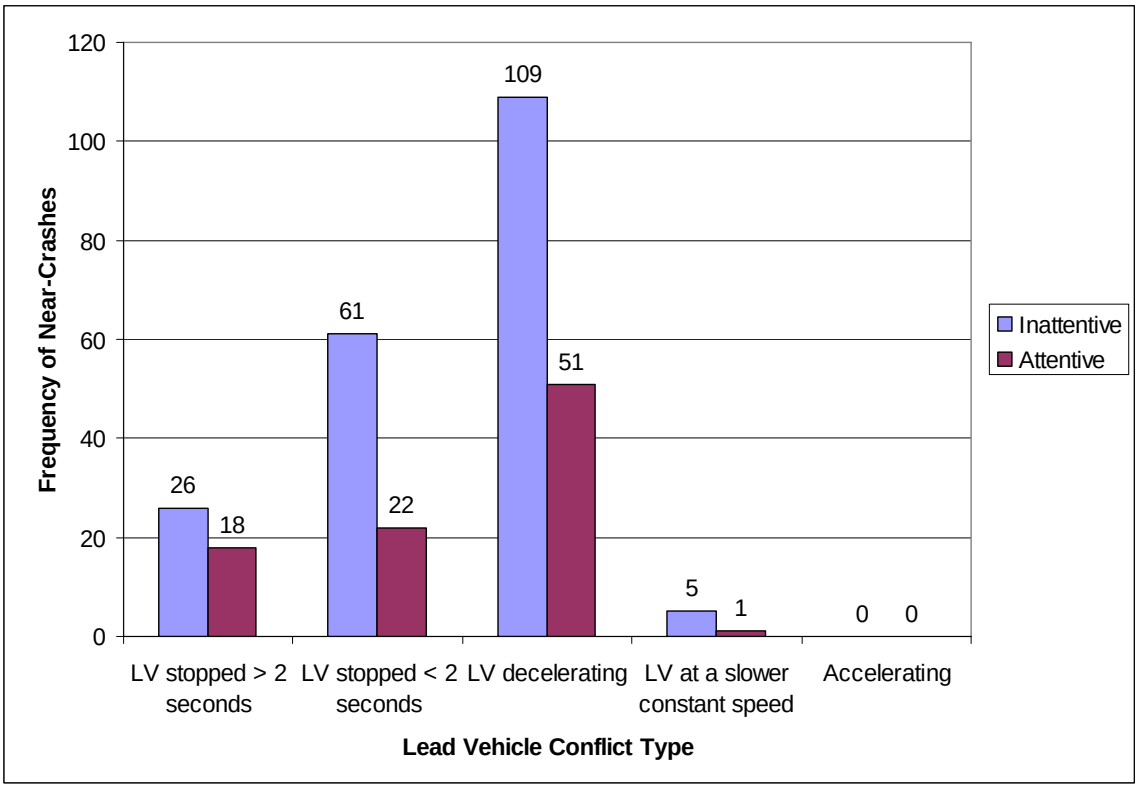


Figure 11.7. Frequency of near-crashes by driver attention level and LV scenario.

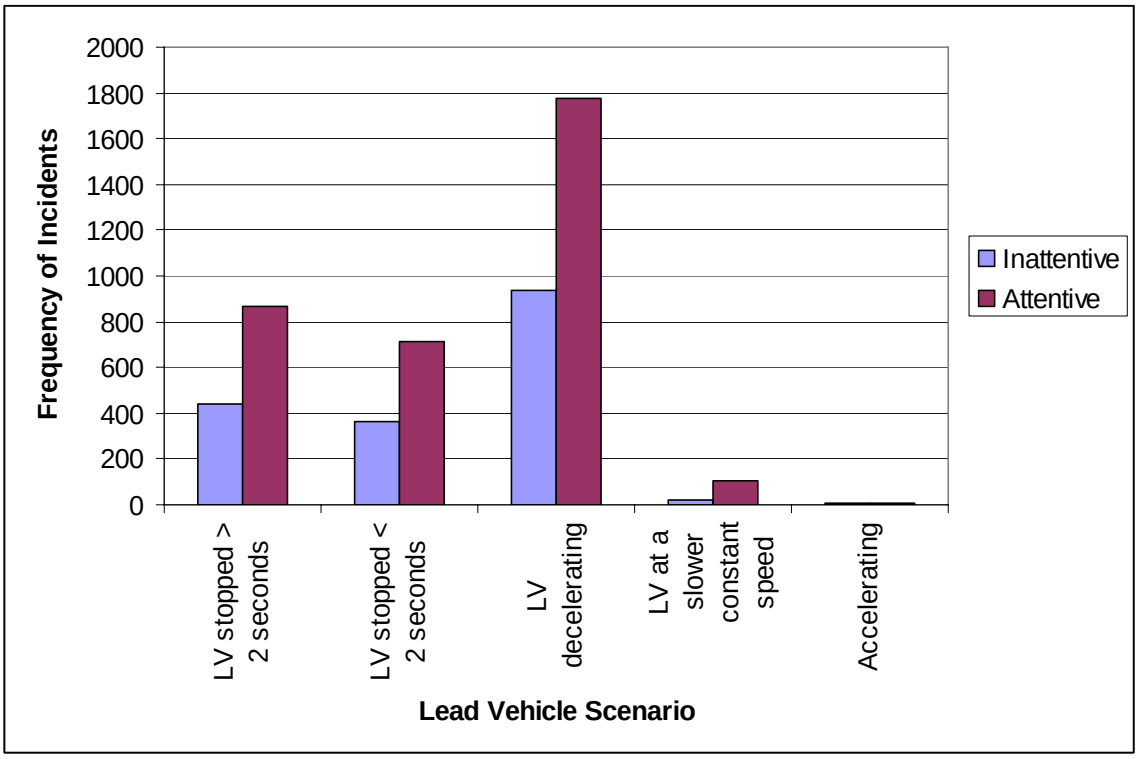


Figure 11.8. Frequency of incidents by driver attention level and LV scenario.

The previous two figures indicate that the most *events overall* occurred in the *lead-vehicle decelerating* case, however, no crashes occurred for this scenario. All of the lead-vehicle conflict crashes were categorized as either *LV stopped* > *2 s* or *LV stopped* ≤ *2 s*. In the *LV stopped* ≤ *2 s* case, this suggests that the driver's expectation regarding LV behavior in the near future may have been violated. That is, drivers might assume that a moving LV is unlikely to rapidly decelerate to a stop (expectation based on prior experience). In some of the *LV stopped* > *2 s* crashes, an inopportune glance away from the roadway appeared to significantly delay the processing of and response to the rapid closure rate. This may be due to an interruption in the continuity of the "looming" cue that occurs when approaching a stopped lead vehicle. These hypotheses should be tested further, although the data clearly suggest that such analyses would be appropriate.

Question 2. For each RE-Lead-vehicle scenario event involving inattention, what is the frequency of each inattentive behavior?

The purpose of this question is to understand how various sources of inattention contribute to events in each of the lead-vehicle kinematic scenarios described above. Please note that for all of the following figures, the number of events with any form of inattention listed as a contributing factor was counted as often as it was deemed a contributing factor. Therefore the emphasis is looking at the frequency of which particular types of inattention were present (e.g., drowsiness) rather than the total number of inattention-related crashes, near-crashes, or incidents. For a complete description of the secondary tasks and all subcategories, please refer to Appendix D.

Lead-vehicle Conflict Overview

Figure 11.9 shows the frequency of each source of inattention for all secondary task categories. This allows comparison of the actual contribution of each of these sources of inattention to lead-vehicle conflicts. Wireless devices (primarily cell phones, but including a few PDAs) were the most frequent contributing factors for lead-vehicle events, followed by passenger-related secondary tasks.

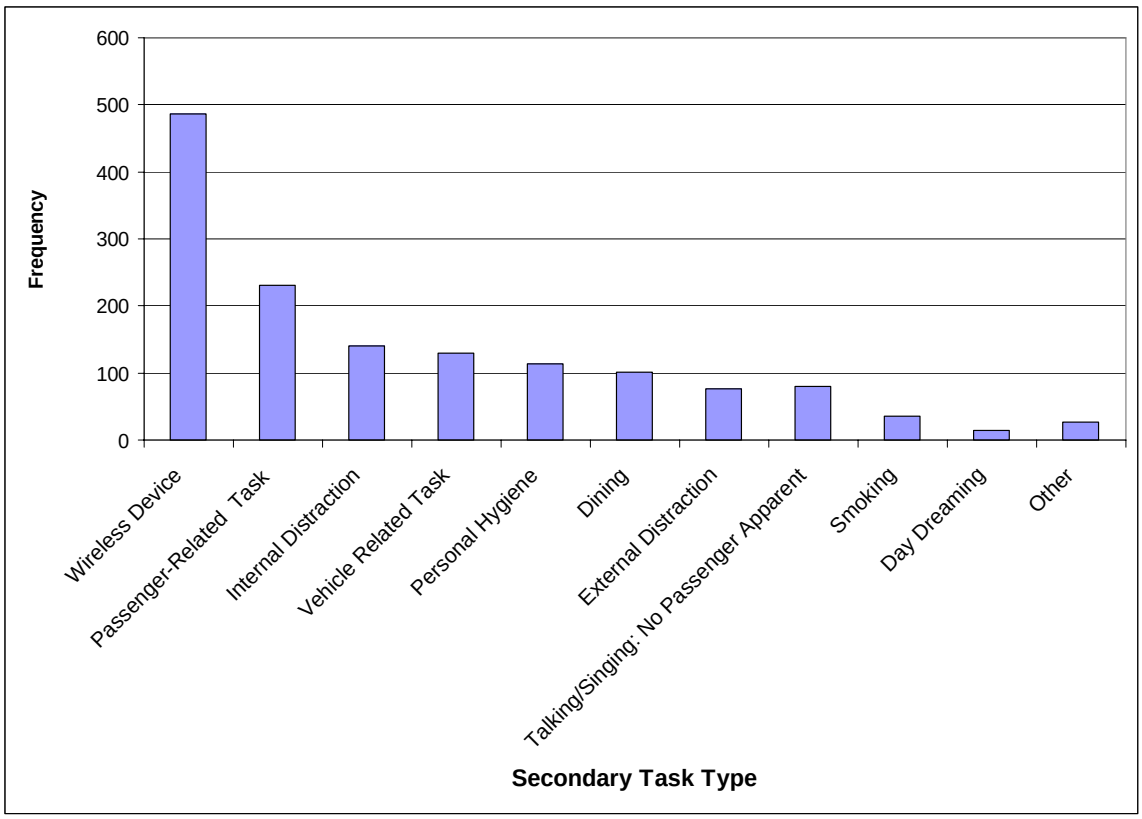


Figure 11.9. Total frequency of secondary task inattention sources for lead-vehicle events.

Figure 11.10 shows the frequency of each secondary task by severity. Note that more lead-vehicle crashes occurred while drivers were engaged in *internal distraction*s and *dining,* whereas many more lead-vehicle near-crashes and incidents occurred while the drivers were engaging in using *wireless devices* or *passenger-related tasks*.

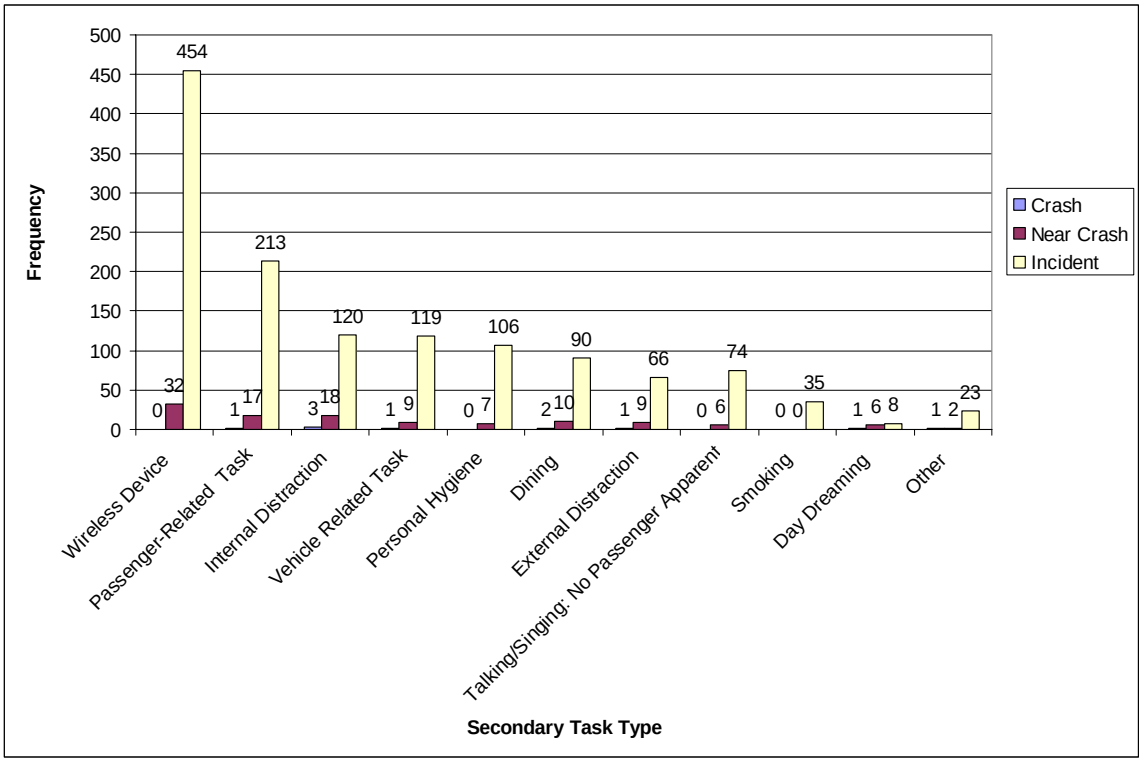


Figure 11.10. Total frequency of secondary task type by severity.

Figure 11.11 shows the contribution of drowsiness to lead-vehicle crashes, near-crashes, and incidents. Note that only one lead-vehicle crash and 33 lead-vehicle near-crashes had drowsiness listed as a contributing factor.

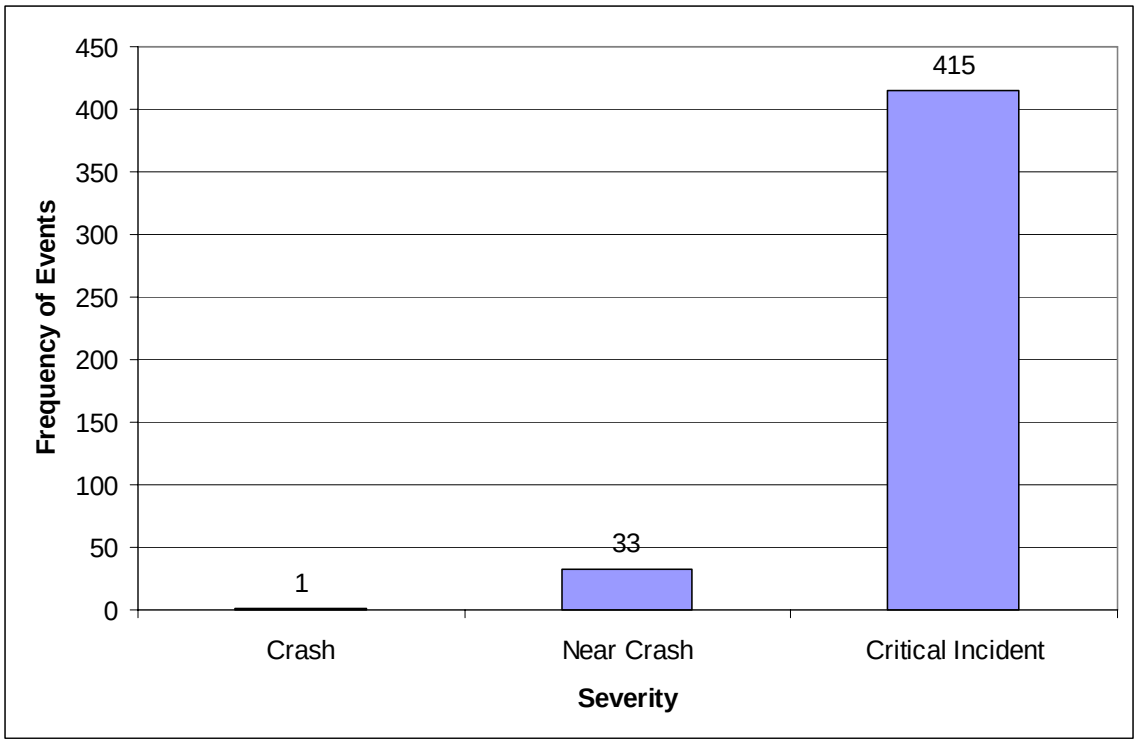


Figure 11.11. Contribution of drowsiness to the overall number of inattention-related lead-vehicle crashes, near-crashes, and incidents.

Figure 11.12 shows the location of drivers' glances for the inattention category, *driving-related inattention to the forward roadway*. For all of the lead-vehicle crashes (3) in this inattention category, drivers were looking out of the left window, thus degrading their peripheral view of the forward roadway. Left window was also the most common category for incidents, however, center mirror glances were most frequent category for near-crashes. Again, the drivers' peripheral view of the forward roadway would be degraded most often by the right and left window glances. This indicates a slight tendency for an increasing event severity level as the glance location becomes more off-roadway-center.

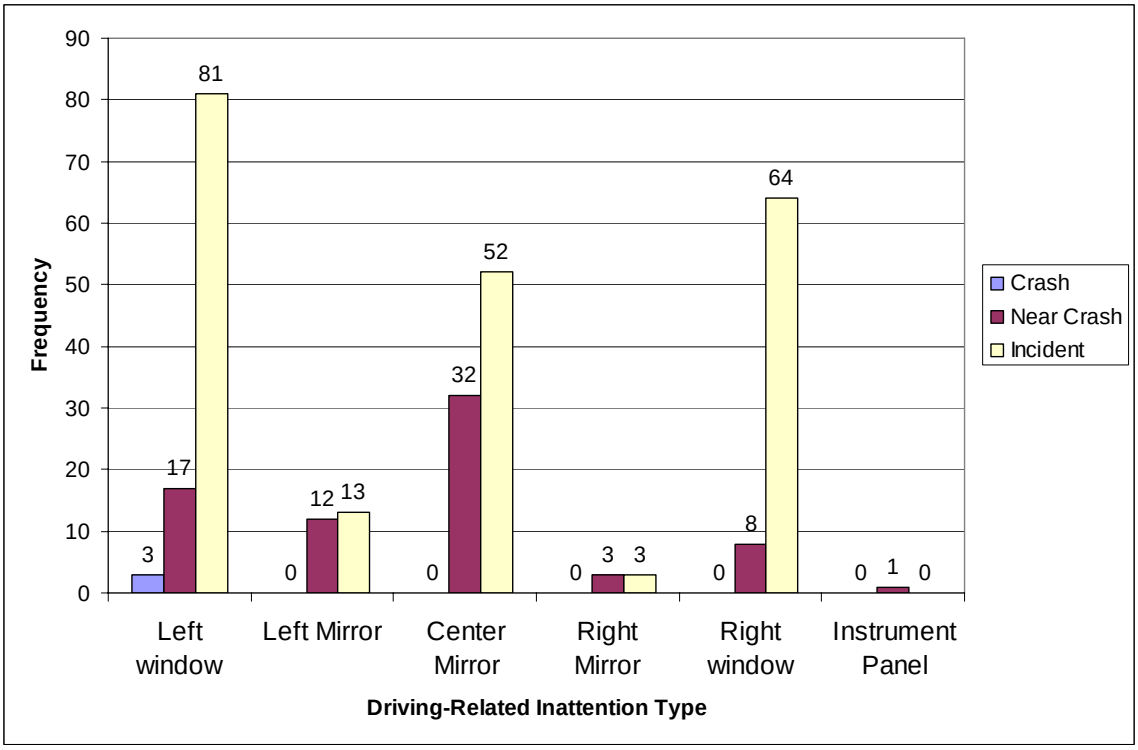


Figure 11.12. Frequency of glance locations for lead-vehicle events categorized as driving-related inattention.

Figure 11.13 presents the crash and near-crash events for nonspecific glances away from the forward roadway. The frequencies presented below represent 23 percent of all inattention-related lead-vehicle crashes and 21 percent of all inattention-related near-crashes. Most nonspecific eyeglances are at unidentified objects internal to the vehicle. Internal distractions were also the most frequent type of secondary inattention for crashes. This finding may indicate that an eyeglance into the vehicle may reduce peripheral vision of the forward roadway, thus leading to more crashes.

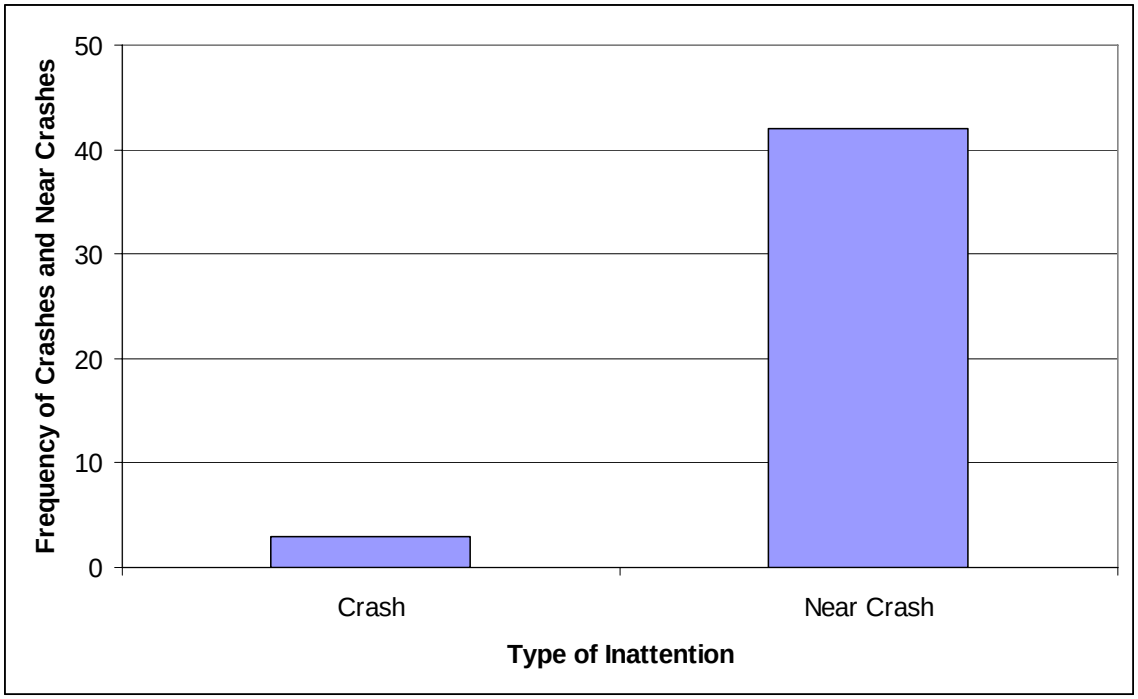


Figure 11.13. Frequency of crash and near-crash events for which a glance to a nonspecific location was present.

Analyses of Inattention-Related Events by Lead-vehicle Kinematic Scenario
In these analyses, four figures will be presented for each lead-vehicle kinematic scenario to better understand how inattention impacts each. The initial figure will show the frequency contributions of each inattention category or combined category for crashes and near-crashes. A second figure will present the frequency of each inattention category for incidents. The final figure will break down the secondary task category (the most frequent type of inattention) into the high level secondary task categories. For this final figure, some events are represented more than once, as multiple types of secondary tasks occurred in some events.

LV Stopped Greater Than 2 s. For this scenario, secondary tasks are the most common source of inattention by a 2 to 1 margin for crashes and a 6 to 1 margin for near-crashes, (Figure 11.14). For incidents, secondary tasks are again more frequent than drowsiness by a margin of 3 to 1 and more frequent than driving-related inattention by 6 to 1 (Figure 11.15).

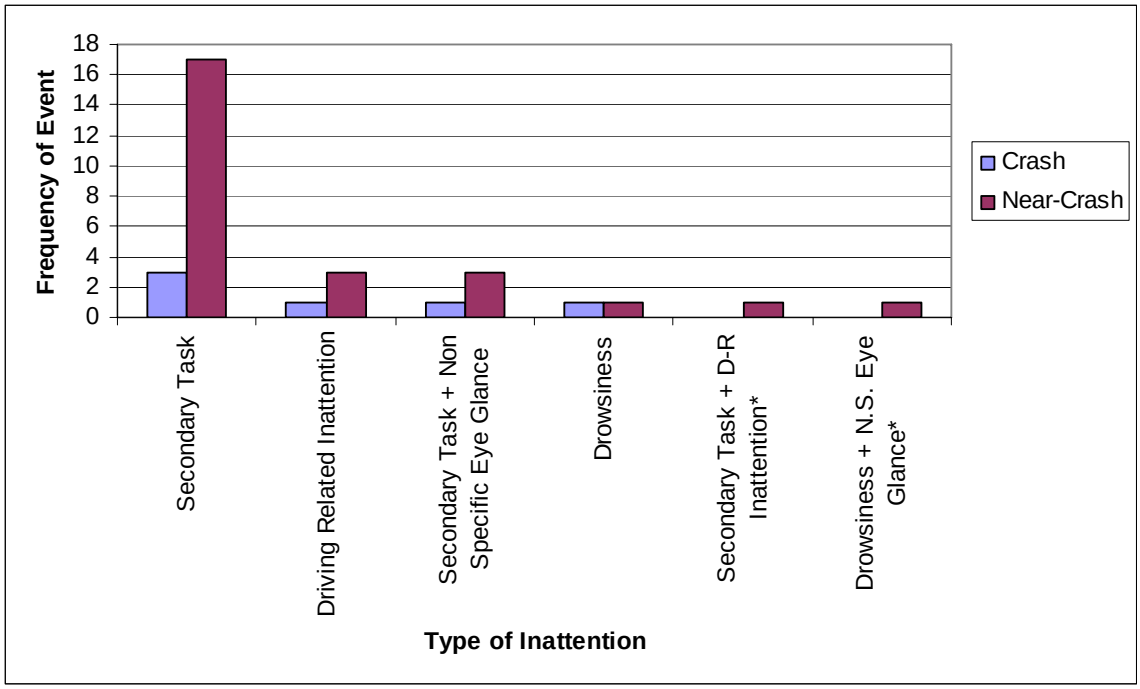


Figure 11.14. Frequency of *LV stopped >2 s* inattention-related events by inattention source and event severity.

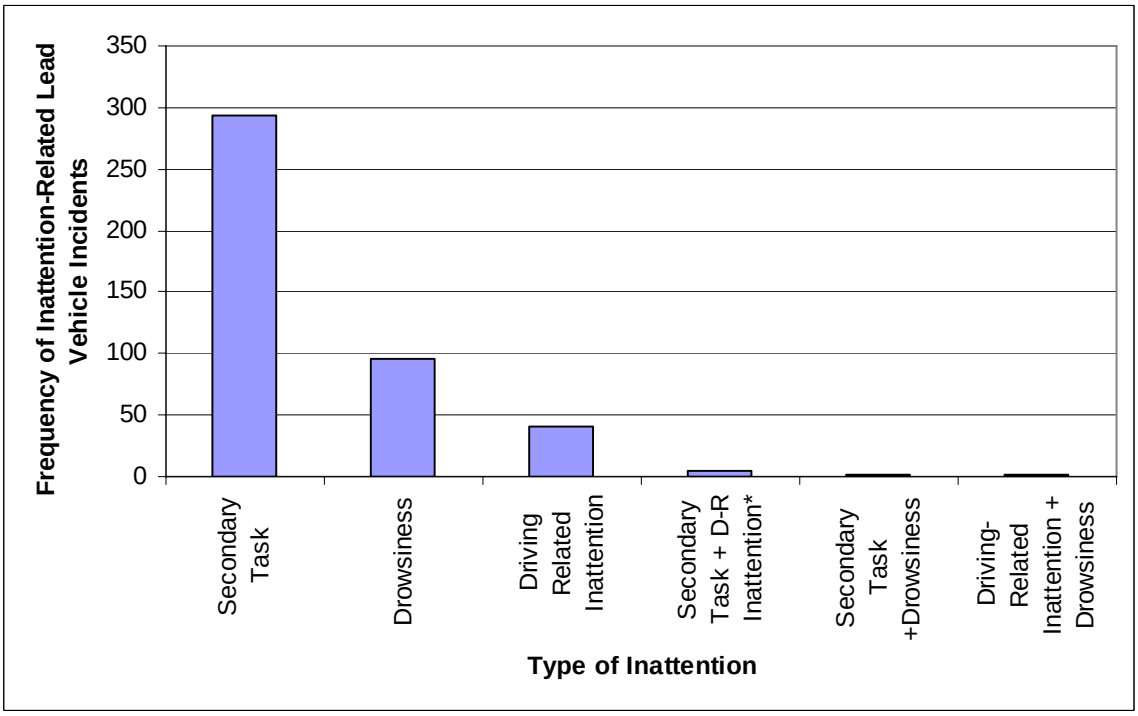


Figure 11.15. Frequency of *LV stopped >2 s* inattention-related incidents by inattention.

Figure 11.16 presents the frequencies for each of the secondary task sources of inattention for crashes, near-crashes, and incidents. If a single crash, near-crash, or incident contained multiple secondary tasks as contributing factors, all sources are presented in the figure below. The types of secondary tasks that were contributing factors to the crashes in this scenario are distributed over *internal distraction, dining, and other*. *Internal distraction* and *wireless devices* were the most frequent sources of secondary task inattention for near-crashes. *Wireless device tasks* occurred more frequently than other types of distractions by over 2 to 1 for incidents. *Passenger-related tasks* and *internal distraction* tasks were also large contributors to incidents.

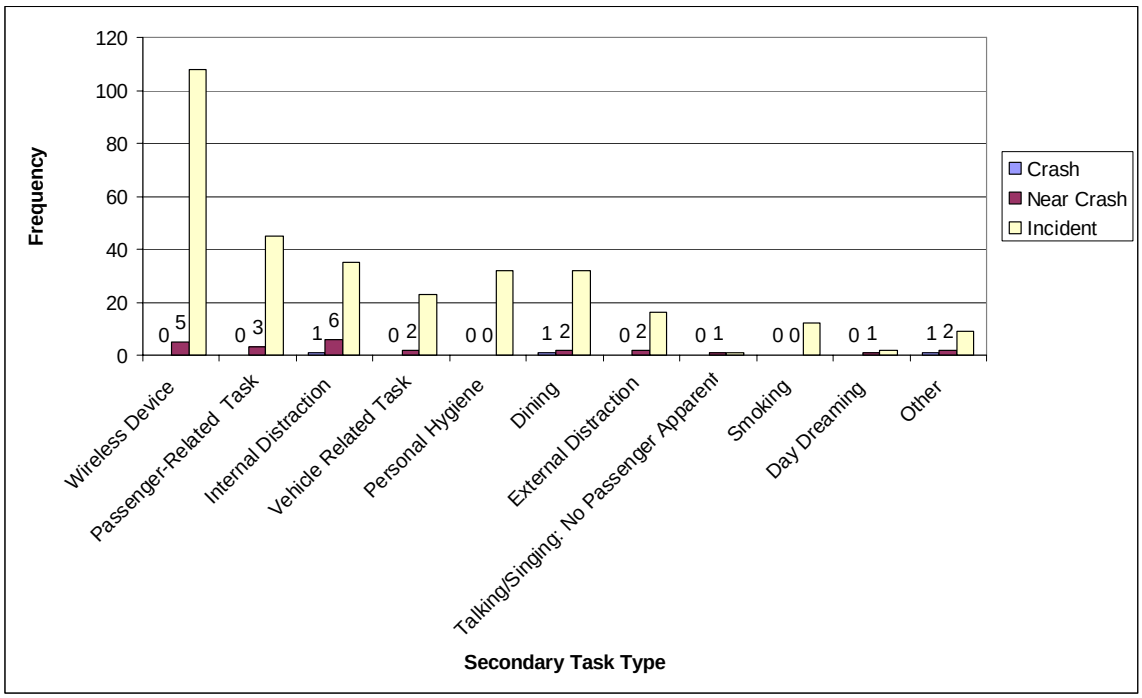


Figure 11.16. Frequency of secondary task inattention sources for *LV stopped >2 s* events by event severity.

LV stopped less than 2 seconds. The inattention-related events that fell under this kinematic scenario were also due primarily to secondary tasks (Figure 11.17). This was true for crashes, near-crashes, and incidents. Note also that *driving-related inattention* and *secondary task + non specific eyeglance* accounted for nearly 66 percent of all inattention-related *lead-vehicle stopped <2 s*. This may suggest that, again, the driver's eyes being off the forward roadway was a significant contributing factor to these crashes. *Drowsiness* and *driving-related inattention to the forward roadway* were the next most frequent categories for near-crashes. For incidents, *drowsiness* contributed to more incidents than did *driving-related inattention to the forward roadway* (Figure 11.18).

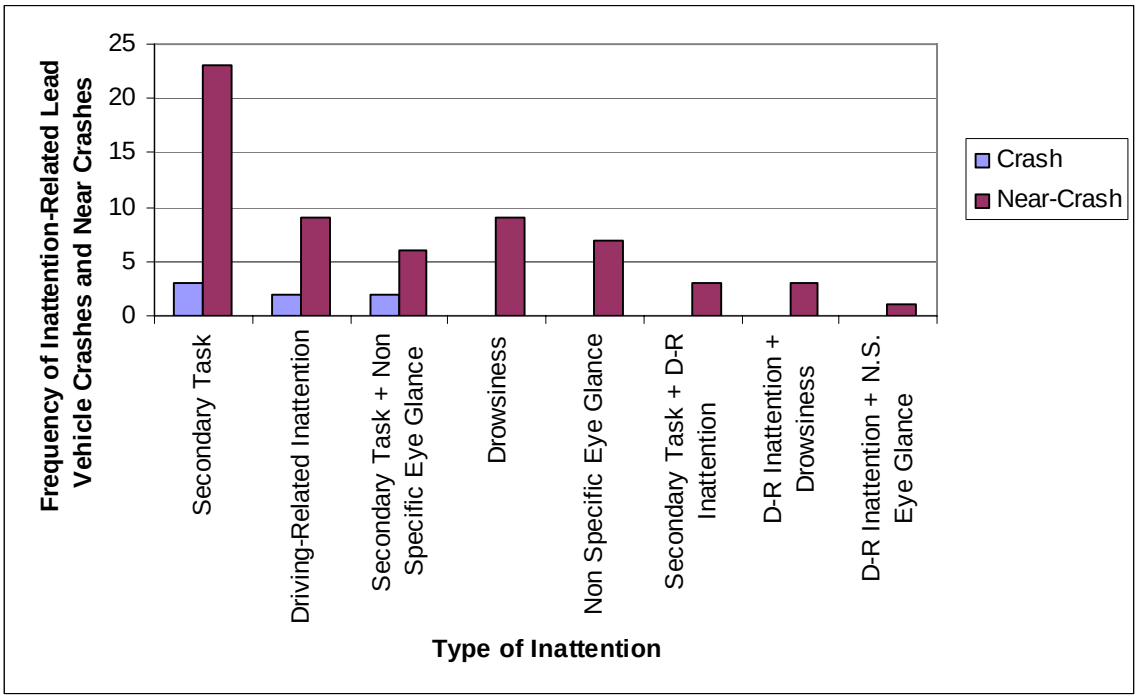


Figure 11.17. Frequency of *LV stopped* < *2 s* inattention-related crashes and near-crashes by event severity.

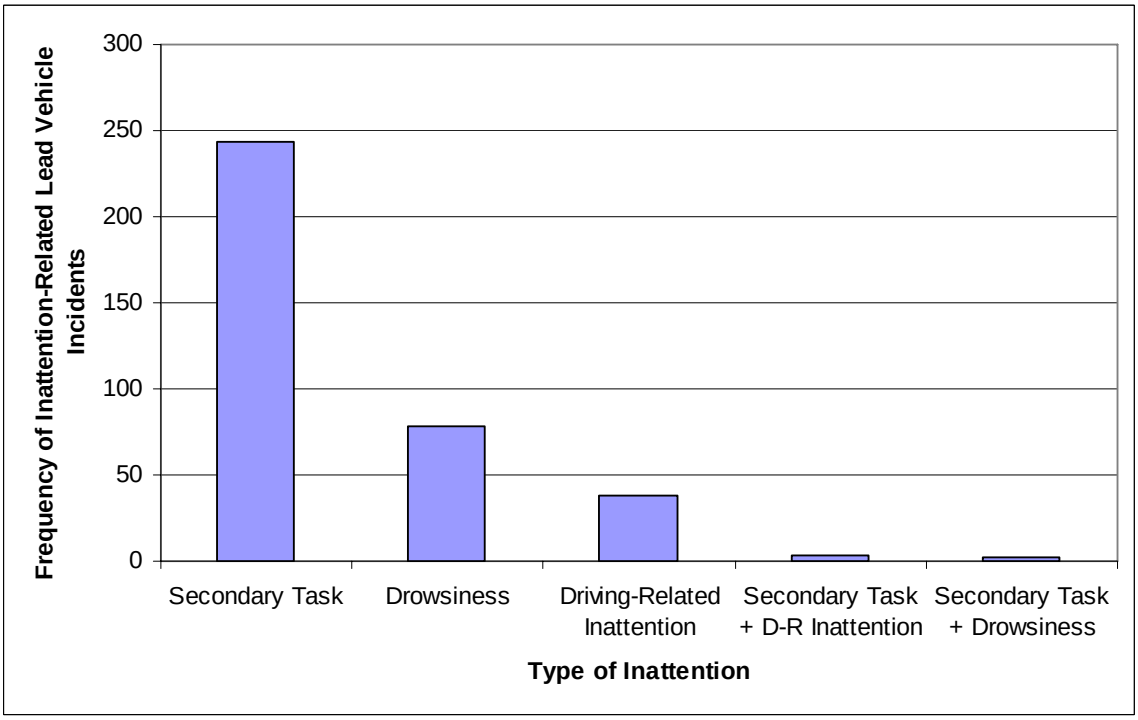


Figure 11.18. Frequency of *LV stopped* < *2 s* inattention-related incidents by inattention source.

For the secondary task types, *wireless device* use was once again the most frequent source of inattention for incidents by a factor of 2 over the next most frequent category of *passenger-*

related inattention and *internal distractions* (Figure 11.19). For near-crashes, *wireless devices*, *internal distractions,* and *passenger-related distractions* were the most frequent sources of inattention. For crashes, the types of secondary distraction were equally spread across 5 different sources with one case each.

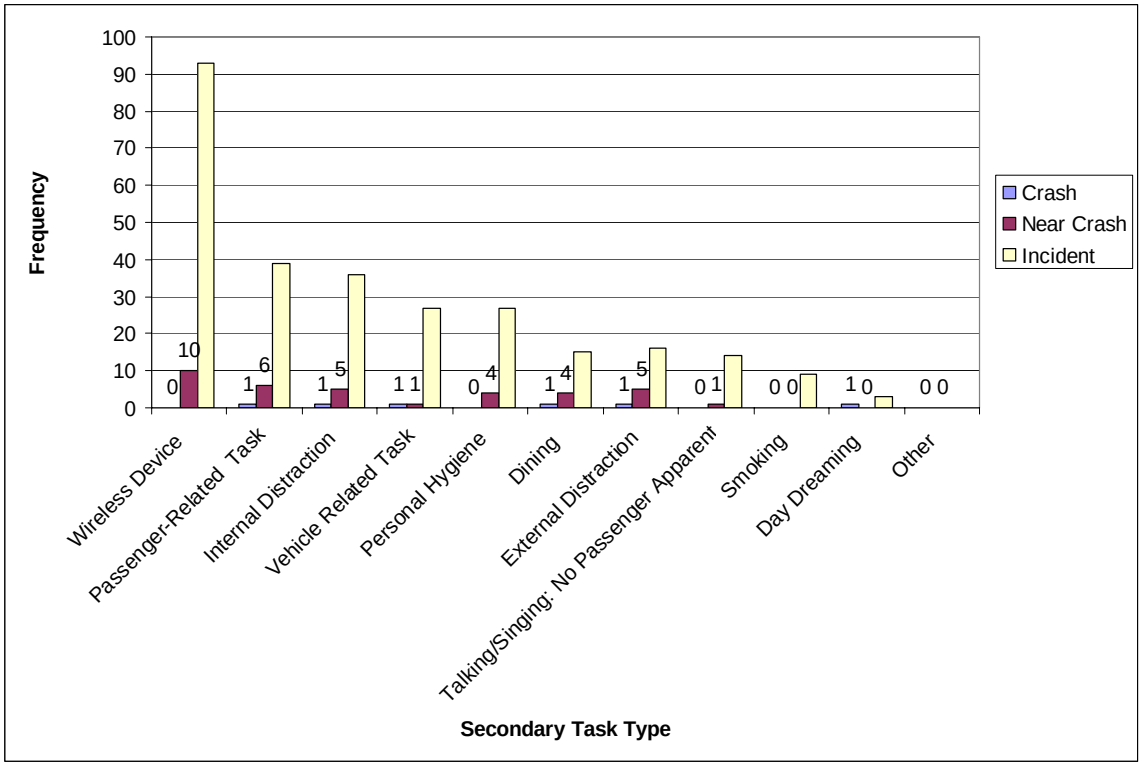


Figure 11.19. Frequency of secondary task inattention sources for *LV stopped* < *2 s* events by event severity.

LV decelerating. This lead-vehicle scenario constituted a majority of the conflicts, however no crashes occurred for this kinematic scenario (Figures 11.20 and 11.21). For near-crashes, secondary tasks were still most common, but by a smaller margin than for incidents, and the second most common category for near-crashes was driving-related inattention. For incidents, secondary tasks outnumbered drowsiness by over 3 to 1 and for near crashes, outnumbered drowsiness by a factor of 4 to 1. Figure 11.22 shows the frequency of secondary tasks for this kinematic scenario. For incidents, *wireless device use* is the most frequent source of secondary task inattention by about 3: 1. For near-crashes, the most frequent source of secondary distraction was *passenger-related inattention*.

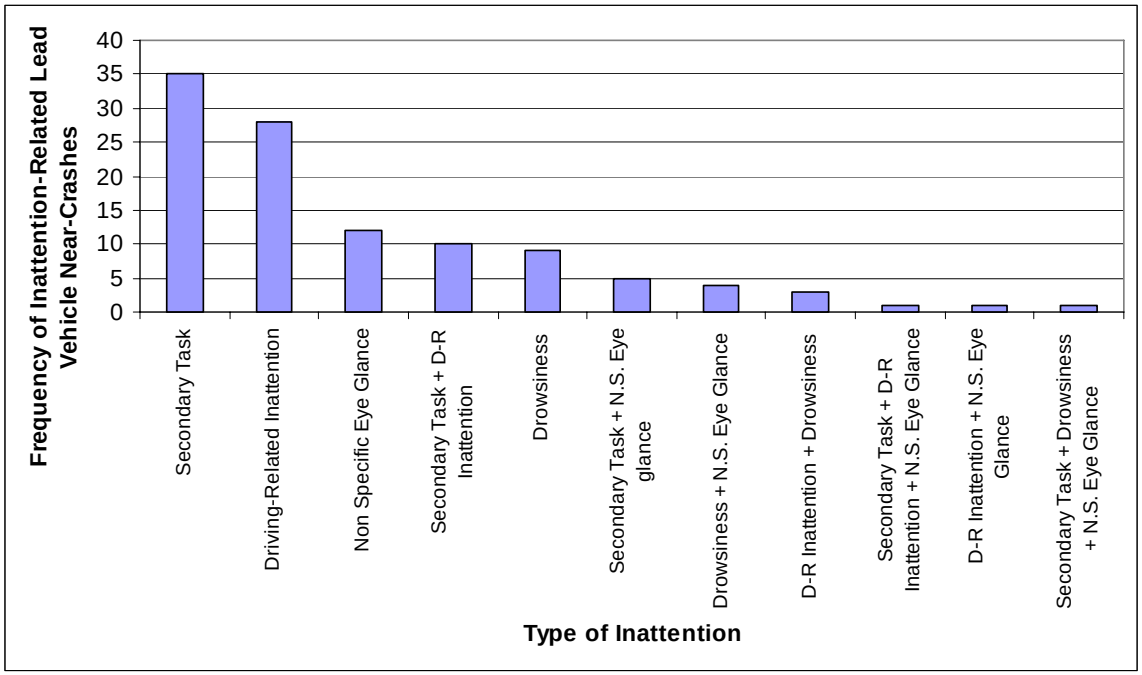


Figure 11.20. Frequency of *LV decelerating* inattention-related near-crashes. Note: There were no crashes in this category.

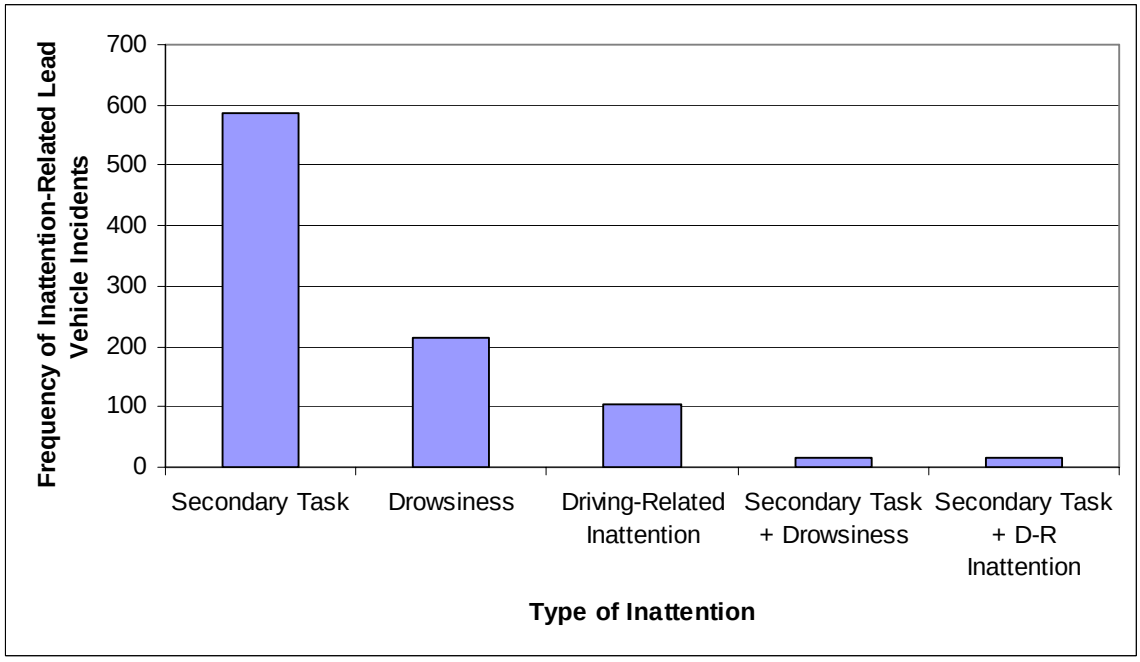


Figure 11.21. Frequency of *LV decelerating inattention*-related incidents.

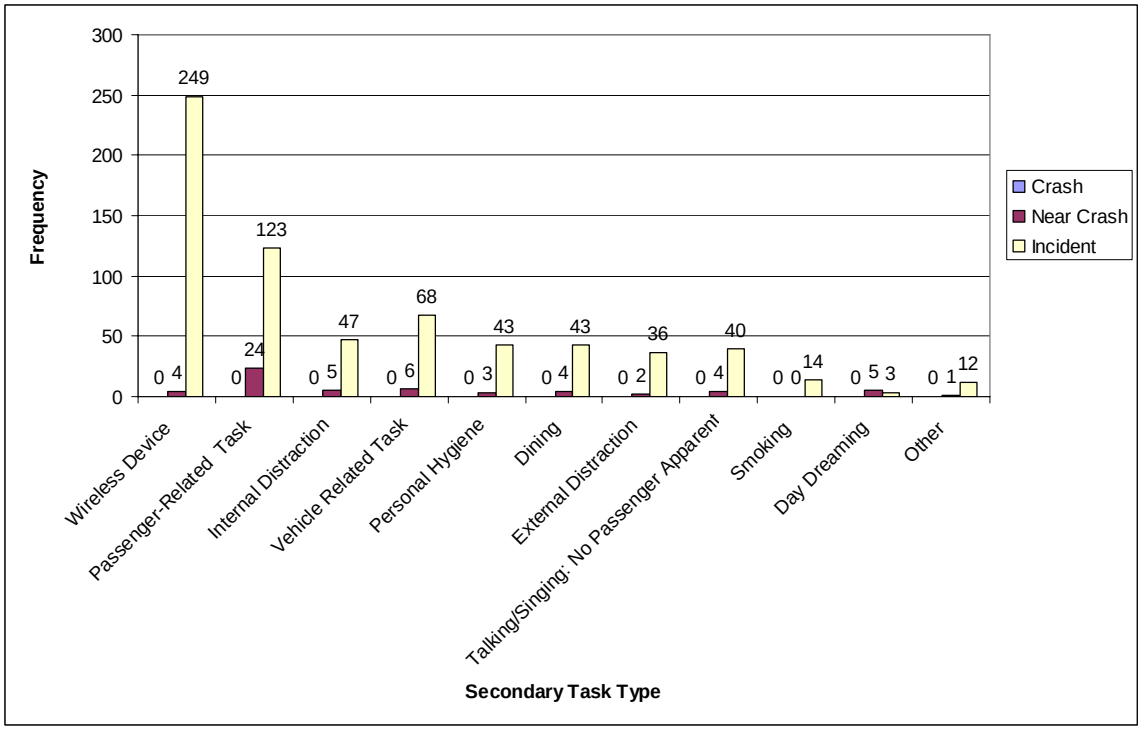


Figure 11.22. Frequency of secondary task inattention sources for *LV decelerating* events by event severity.

<u>LV moving at a slower constant speed.</u> *Secondary tasks* again contributed to more incidents for this scenario than did *drowsiness* for incidents (Figures 11.23 and 11.24). However, the most frequent source for near-crashes was *driving-related inattention to the forward roadway* (although there were only 5 near-crash events for this scenario). The secondary task distribution presented in Figure 11.25 shows that cell phone use contributed to the only near-crash that occurred for this category. For incidents, passenger-related tasks outnumbered wireless devices and personal hygiene.

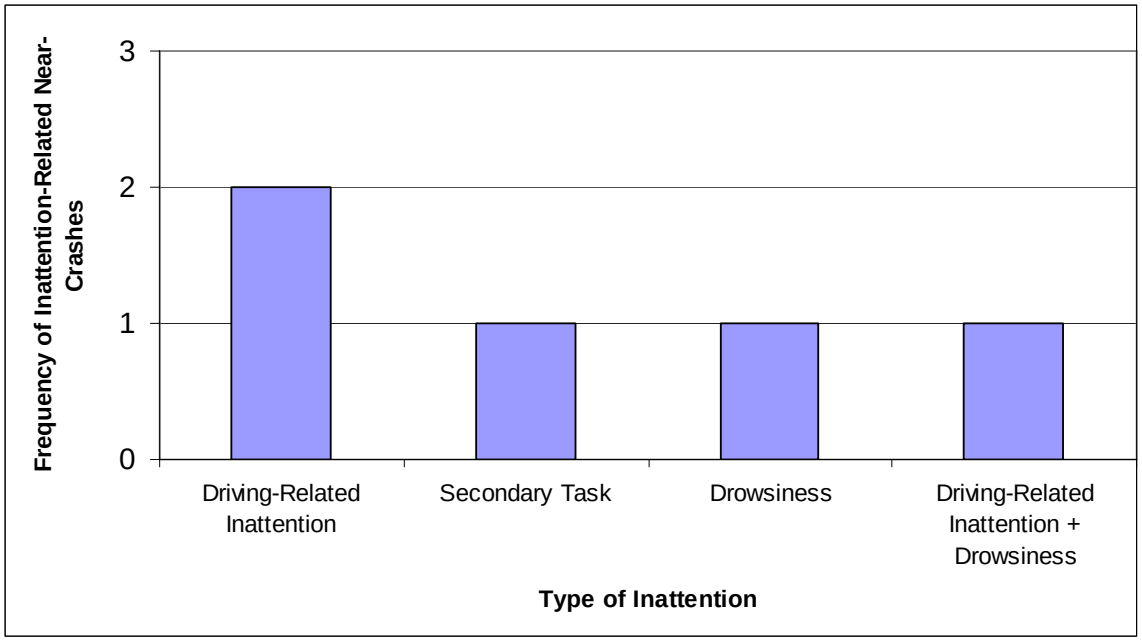


Figure 11.23. Frequency of *LV moving at slower constant speed* inattention-related near-crashes. Note: There were no crashes in this category.

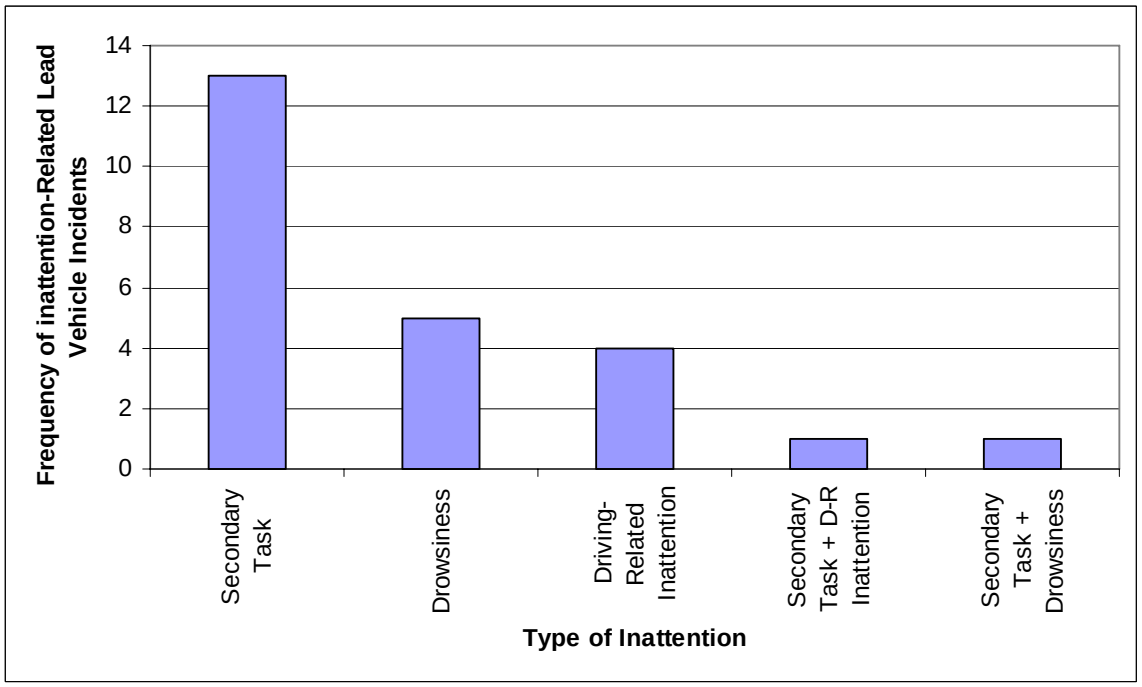


Figure 11.24. Frequency of *LV traveling at a slower, constant speed inattention*-related incidents.

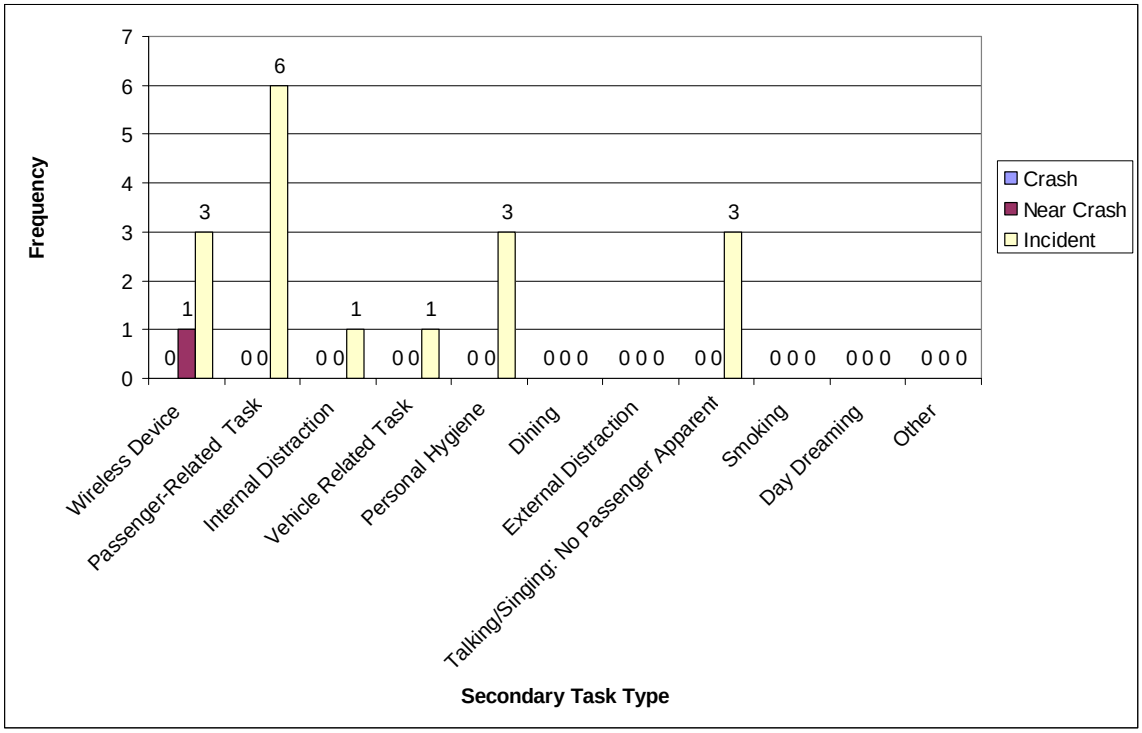


Figure 11.25. Frequency of secondary task inattention sources for *LV moving at slower constant speed* events by event severity.

Lead-vehicle accelerating. There were only 3 cases of lead-vehicle conflicts in which the lead vehicle was accelerating and the SV driver was engaged in an inattention-related task. Each of these events involved a different distraction task: wireless device, personal hygiene, and internal distraction. No drowsiness or driving-related inattention sources were identified as contributing factors for this type of lead-vehicle conflict.

Analysis of the Most Frequent Secondary Tasks

Wireless device analysis. There were no lead-vehicle crashes attributed to wireless device tasks; however, a large percentage of the near-crashes and incidents involved wireless devices, especially cell phones. As shown in Figure 11.26, wireless device use appeared to affect more of the incidents and near-crashes in which the lead vehicle was stopped or decelerating, suggesting that drivers may have experienced difficulty judging closing rates to the lead vehicle while using these devices.

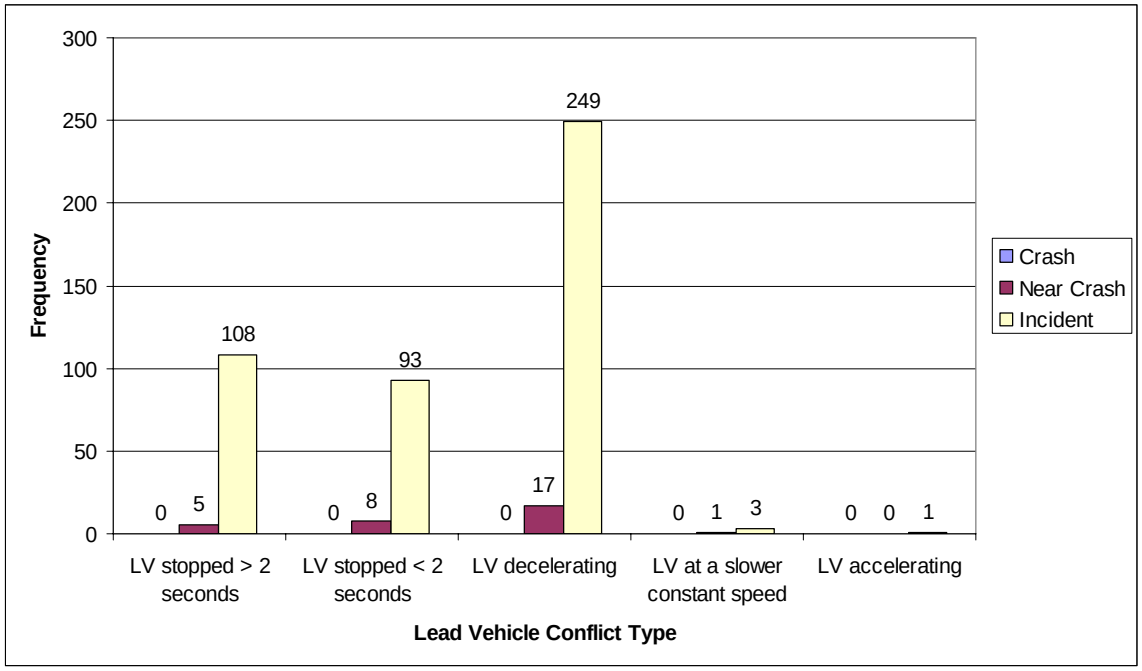


Figure 11.26. Frequency of LV events for wireless devices by scenario and severity.

Passenger-related inattention analysis. Only one crash was attributed to the passenger-related inattention secondary task category (Figure 11.27). More of the near-crashes and incidents occurred for the decelerating and stopping categories, which again suggests that drivers may have experienced difficulty in judging time-to-collision while engaging in conversation.

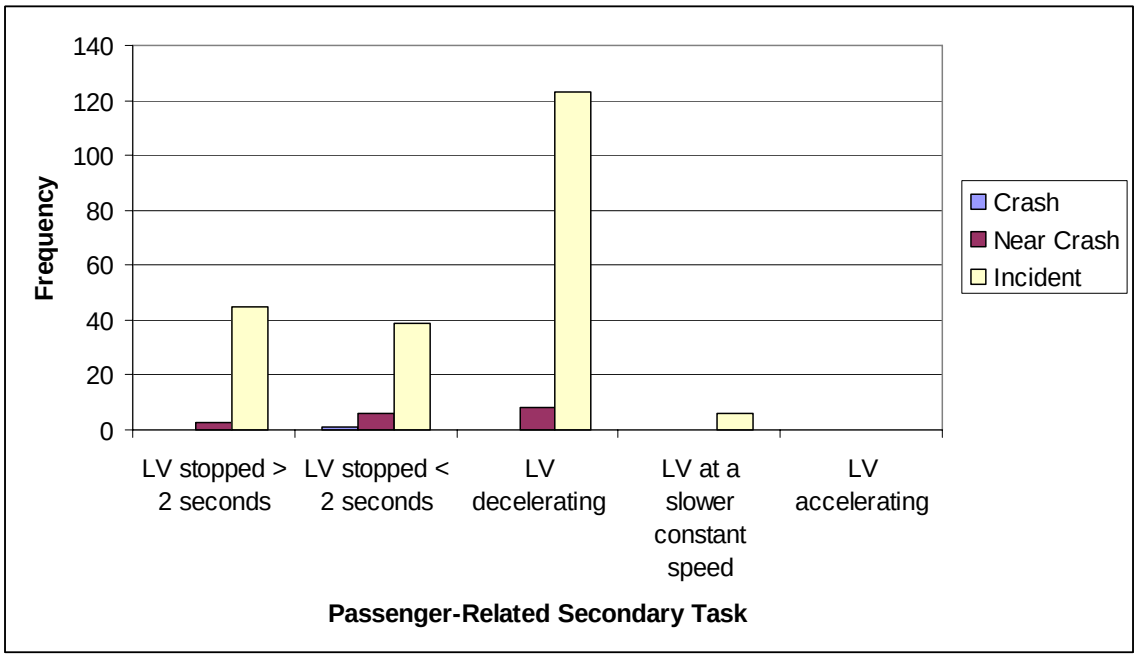


Figure 11.27. Frequency of LV events for passenger-related inattention by scenario and severity.

DISCUSSION

One of the most significant findings of the Chapter 11, *Goal 7* analyses is that 93 percent of all lead-vehicle crashes occurred when the driver was inattentive. All crashes occurred when the *lead vehicle was stopped greater than 2 seconds* or *lead vehicle was stopped less than or equal to 2 s*, however, *lead-vehicle decelerating* was the most frequently occurring kinematic scenario for near-crashes and incidents.

Of particular interest was that glances away from the forward roadway contributed to 4 of the 5 most frequent types of inattention for both crashes and near-crashes. An important finding of this research is that inopportune eyeglances (those close in time to the precipitating factor) are a significant contributing factor to crashes and near-crashes.

Secondary tasks generally contributed to the greatest number of lead-vehicle events (65%), while drowsiness contributed to 22 percent and driving-related inattention contributed to 12 percent of these events. Cell phone operations and passenger-related distractions were the two most frequently occurring secondary task sources. While cell phone use was a much more frequent contributor to incidents and near-crashes than any other secondary task, it did not contribute to any lead-vehicle conflict crashes. However, cell phone use did contribute to several crashes of other types, as reported in other chapters of this report.

CHAPTER 12: GOAL 8, CHARACTERIZE EACH OF THE 4 RE SCENARIOS IN RELATION TO HEINRICH'S TRIANGLE

BACKGROUND

Many industrial safety researchers face similar challenges to those experienced by transportation researchers when attempting to obtain a measure of safety or predict the probability that an accident will occur given certain circumstances. In most settings, accidents leading to injury or death are rare events. Therefore, a paradox is present in that any method that relies on the collection of sufficient accident (or crash) data is reactive and not proactive. Alternatively, any method that relies on indirect indicators of accident (or crash) risk is often not predictive of accidents (or crashes).

Heinrich, Petersen, and Roos (1980) developed a hazard analysis technique based on the underlying premise that for every injury accident, there were many similar accidents in which no injury occurred. Variations of the hazard analysis technique have been developed for use in transportation research and specifically on instrumented vehicles. This modification involves cameras being strategically placed on a vehicle to determine the number of safety-related events (crash, near-crash, or incident) for a particular driver (Mollenhauer, 1998; Hanowski et al., 2000; Dingus et al., 2001; Wierwille et al., 2002). Hanowski et al. (2000) and Dingus et al. (2001) used this modified version of the hazard analysis technique by videotaping commercial vehicle drivers and the surrounding driving environment to identify incidents, near-crashes, and crashes. This technique proved useful in identifying the effects of drowsiness in truck driving and in determining the prevalence of drowsiness among these commercial truck drivers.

For the 100-Car Study, cameras placed in 100 vehicles ultimately recorded 241 drivers during the course of one year. Previous studies have used this technique, but did not collect sufficient crash data to investigate the relationships between incidents, near-crashes, and crashes. An argument can be made that the current study did not collect enough crash data to make relevant comparisons between crashes, near-crashes, and incidents for all event types, but it does present a landmark study in that events of each type were present from the same drivers in the same database. Moreover, for rear-end crashes, a sufficient number of crashes, near-crashes, and incidents were present such that comparisons can be discussed.

DATA ANALYSIS OVERVIEW

This analysis was performed for subject vehicles in the striking vehicle role for conflict with lead-vehicle events. It should be noted that although 15 crash events will be used throughout this analysis, information on 16 events was recorded. In one case, a data system failure prevented collection of relevant information. The driver did report the event and post-event information was collected. However, due to the lack of data regarding precipitating factor, this event was not considered in the following analysis. The frequency counts of the various rear-end conflict types for the three event severity levels were calculated and are discussed in relation to the levels of Heinrich's Triangle.

Frequency counts were also modeled using the Poisson distribution, which has been used to model rare events (i.e., crashes). This distribution is often employed to model cases in which a

number of events occur in a unit of density. For example, the number of crashes can be modeled per MVMT. The Poisson distribution has a single parameter, often called the intensity or rate, which was estimated for the RE scenarios. This was accomplished using the number of events and the estimates for vehicle miles traveled. In addition, estimates of variability and 95 percent confidence intervals were calculated for each of these parameter estimates. Following Tijerina (2004), the 95 percent confidence intervals were based on large sample normal approximations. These confidence intervals are further used to determine the relevance of the near-crash to crash ratios.

Data Included in the Analysis

Lead-vehicle conflict data were used for all 241 drivers who participated in the study and rates were calculated based on the mileage collected in the 100-Car Study.

It may be noted that the exposure for MVMT is different for crashes than for near-crashes and incidents. In the study, 1.84 million vehicle miles were calculated based on odometer readings. Due to various component failures, it was not possible to determine the incidents and near-crashes for 0.43 MVMT, leaving 1.37 MVMT. Nevertheless, when crashes occurred, drivers reported them, thereby allowing use of the 1.84 MVMT.

Note that analyses were not conducted for following-vehicle events because there were far fewer following-vehicle events as compared to lead-vehicle events. One reason was the difference in the radar signatures for a forward versus rear-facing radar. Essentially, forward-facing radar has many more objects to discern since any static object being approached represents a potential threat. Alternatively, rear-facing radar only needs to produce a signature for objects moving toward the vehicle since all other targets are increasing in range as the vehicle moves forward. A second reason was that video confirmation of following-vehicle events was difficult due to rear camera placement and angle. Simply put, the rear-facing camera did not create a natural viewpoint from which to judge following events.

Question 1. What is the relative frequency of each RE lead-vehicle scenario resulting in a crash, near-crash, or incident?

Heinrich's Triangle for All Lead-vehicle Conflict Types

Figure 12.1 shows the Heinrich's Triangle for all lead-vehicle conflict types. The ratio of incidents to near-crashes is approximately 18:1 and approximately 18:1 for near-crashes to crashes.

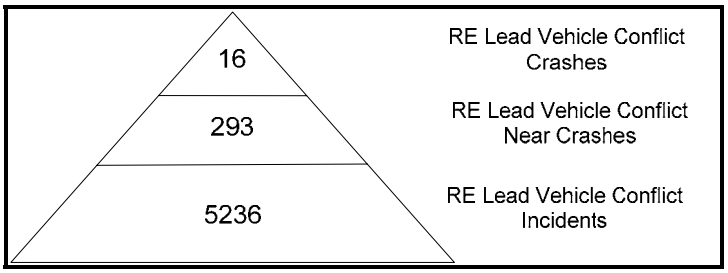


Figure 12.1. Heinrich's Triangle showing the relative occurrence of lead-vehicle conflict events by severity.

The estimates of the Poisson rate per MVMT, variance, and 95 percent confidence intervals are shown in Table 12.1. Each type of RE conflict will be discussed, and comparisons will be made to the overall ratios and rate.

Table 12.1. Lead-vehicle conflicts modeled using Poisson distribution to determine event rate per MVMT, variance, and 95 percent confidence intervals.

	Count	Exposure per MVMT	Rate per MVMT	Variance (rate)	STD (rate)	Lower 95% CI for rate/MVMT	Upper 95% CI for rate/MVMT
RE Crashes	16	1.84	8.7	4.73	2.17	4.43	12.96
RE Near-Crashes	293	1.37	213.87	156.11	12.49	189.38	238.36
RE Incidents	5,236	1.37	3,821.90	2,789.71	52.82	3,718.38	3,925.42

These calculations suggest that lead-vehicle crashes occur at a rate of approximately 9 per MVMT within an approximate confidence interval of 4 to 12. This study observed 16 lead-vehicle crashes, which seems reasonable as the number of vehicle miles traveled is approaching 2 million miles.

Heinrich's Triangles by Lead-vehicle Conflict Type

The following figures show the resulting Heinrich's Triangles for each lead-vehicle scenario. Note that Figure 12.1 represents all of the lead-vehicle conflict types, with a total of 15 crashes. However, one of the crashes had a precipitating factor other that differed from the RE conflicts. Therefore, only 14 crashes were available for the following analyses, and the Heinrich's Triangles that follow do not sum to the Triangle presented above.

<u>LV stopped > 2 s.</u> Reductionists identified approximately 1,305 LV stopped > 2 seconds events (Figure 12.2). The ratios between the two levels were 30:1 for incidents to near-crashes and 6:1 for near-crashes to crashes. It is interesting to note that the ratio drops dramatically as the severity of the event increases, which was not the case for overall lead-vehicle conflicts.

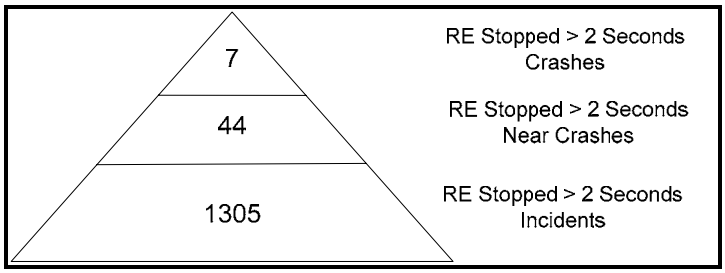


Figure 12.2. Heinrich's Triangle showing the relative occurrence of *LV stopped >2 s* events by severity.

Using a Poisson distribution, the event rate per MVMT, variance, and 95 percent confidence intervals are shown in Table 12.2. *LV stopped >2 s* crashes occur at a rate of approximately 4 per MVMT within a confidence interval of approximately one to 7 crashes per MVMT.

Table 12.2. *LV stopped >2 s* data modeled using Poisson distribution to determine event rate per MVMT, variance, and 95 percent confidence intervals.

	Count	Exposure per MVMT	Rate per MVMT	Variance (rate)	STD (rate)	Lower 95% CI for rate/MVMT	Upper 95% CI for rate/MVMT
RE Crashes	7	1.84	3.80	2.07	1.44	0.99	6.62
RE Near-Crashes	44	1.37	32.12	23.44	4.84	22.63	41.61
RE Incidents	1305	1.37	952.55	695.30	26.37	900.87	1004.24

LV stopped ≤2 s. This scenario type was the third most frequent of the lead-vehicle scenario types and was the only other lead-vehicle scenario with crashes. The ratio of critical incidents to near-crashes is 13:1 and approximately 12:1 for near-crashes to crashes (Figure 12.3). For overall conflicts, the ratios of incidents to near-crashes and of near-crashes to crashes were very similar to those seen for this scenario.

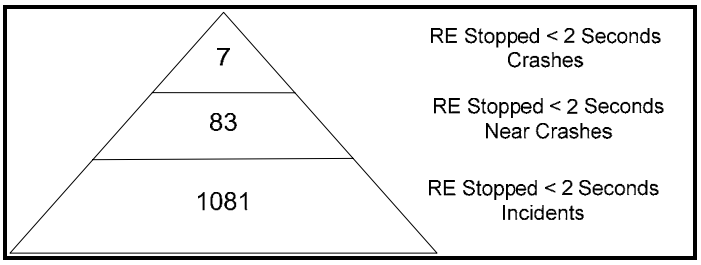


Figure 12.3. Heinrich's Triangle showing the relative occurrence of *LV stopped ≤ 2 s* events by severity.

Using a Poisson distribution, the event rate per MVMT, variance, and 95 percent confidence intervals are shown in Table 12.3. *LV stopped ≤ 2 s* crashes occur at a rate of approximately 4 per MVMT within a confidence interval of approximately 1 to 7. For crashes, this rate is essentially the same as for *LV stopped > 2 s*.

Table 12.3. *LV stopped ≤ 2 s* data modeled using Poisson distribution to determine event rate per MVMT, variance, and 95 percent confidence intervals.

	Count	Exposure per MVMT	Rate per MVMT	Variance (rate)	STD (rate)	Lower 95% CI for rate/MVMT	Upper 95% CI for rate/MVMT
RE Crashes	7	1.84	3.80	2.07	1.44	0.99	6.62
RE Near-Crashes	83	1.37	60.58	44.22	6.65	47.55	73.62
RE Incidents	1081	1.37	789.05	575.95	24.00	742.01	836.09

LV decelerating. While this lead-vehicle scenario type was the second most frequently occurring, no crashes of this type were reported or identified in the database (Figure 12.4). Evidently, while this type of kinematic scenario is a relatively common occurrence, it does not appear to pose as great a crash risk as the *LV stopped* scenarios. The ratio between incidents and near-crashes for the *LV decelerating* events is 17:1 and is fairly similar to the overall lead-vehicle conflict ratio as well as *LV stopped ≤ 2 s*.

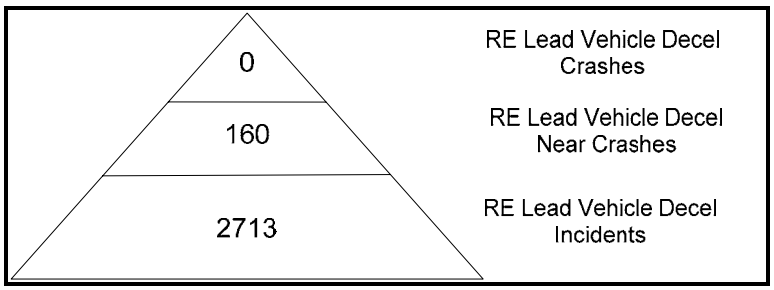


Figure 12.4. Heinrich's Triangle showing the relative occurrence of *LV decelerating* events by severity.

Using a Poisson distribution, the event rate per MVMT, variance, and 95 percent confidence intervals are shown in Table 12.4.

Table 12.4. *LV decelerating* data modeled using Poisson distribution to determine event rate per MVMT, variance, and 95 percent confidence intervals.

	Count	Exposure per MVMT	Rate per MVMT	Variance (rate)	STD (rate)	Lower 95% CI for rate/MVMT	Upper 95% CI for rate/MVMT
RE Crashes	0	1.84	0.00	0.00	-	-	-
RE Near-Crashes	160	1.37	116.79	85.25	9.23	98.69	134.88
RE Incidents	2713	1.37	1980.29	1445.47	38.02	1905.78	2054.81

LV Moving at a Slower Constant Speed. Very few valid events were identified in the database as lead-vehicle conflicts during which the lead vehicle was moving at a slower, constant speed as compared to the subject vehicle (Figure 12.5). None of crashes identified by reductionists or reported by drivers had a slower moving lead vehicle as the precipitating factor. The ratio of incidents to near-crashes (21:1) is slightly higher than the ratio for all lead-vehicle scenarios.

However, it would appear that this kinematic condition is not as common and does not pose a serious crash threat.

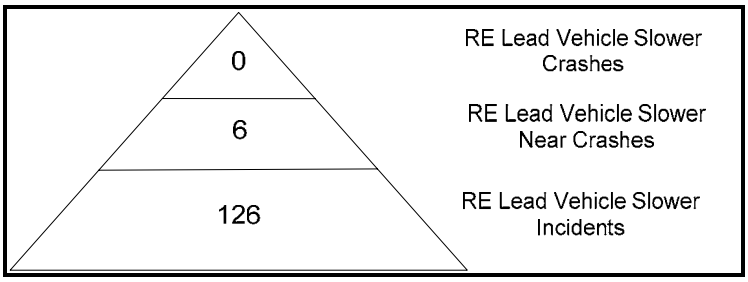


Figure 12.5. Heinrich's Triangle showing the relative occurrence of *LV moving at a slower constant speed* event by severity.

Using a Poisson distribution, the event rate per MVMT, variance, and 95 percent confidence intervals are shown in Table 12.5.

Table 12.5. *LV moving at slower constant speed* data modeled using Poisson distribution to determine event rate per MVMT, variance, and 95 percent confidence intervals.

	Count	Exposure per MVMT	Rate per MVMT	Variance (rate)	STD (rate)	Lower 95% CI for rate/MVMT	Upper 95% CI for rate/MVMT
RE Crashes	0	1.84	0.00	0.00	-	-	-
RE Near-Crashes	6	1.37	4.38	3.20	1.79	0.88	7.88
RE Incidents	126	1.37	91.97	67.13	8.19	75.91	108.03

LV Accelerating. This type of scenario was identified only infrequently by reductionists as resulting in valid events (Figure 12.6). None of the crashes or near-crashes identified by reductionists or reported by drivers had an accelerating lead-vehicle. The low frequency of occurrence, as well as no crash data, suggests that this type of kinematic scenario is uncommon and not a high crash risk for drivers.

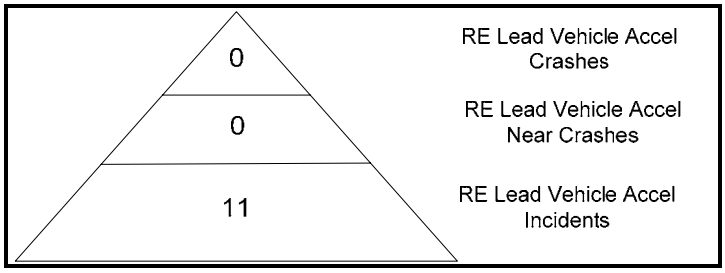


Figure 12.6. Heinrich's Triangle showing the relative occurrence of *LV accelerating* events by severity.

Using a Poisson distribution, the incident rate per MVMT, variance, and 95 percent confidence intervals for incidents are shown in Table 12.6.

Table 12.6. *LV accelerating* data modeled using Poisson distribution to determine event rate per MVMT, variance, and 95 percent confidence intervals

	Count	Exposure per MVMT	Rate per MVMT	Variance (rate)	STD (rate)	Lower 95% CI for rate/MVMT	Upper 95% CI for rate/MVMT
RE Crashes	0	1.84	0.00	0.00	-	-	-
RE Near-Crashes	0	1.37	0.00	0.00	-	-	-
RE Incidents	11	1.37	8.03	5.86	2.42	3.28	12.77

DISCUSSION

When reviewing the frequencies for each lead-vehicle conflict type by severity, an interesting finding is that the frequency of incidents for *LV decelerating* conflicts is roughly twice the frequency for *LV stopped >2 s* and *LV stopped ≤2 s*. Nonetheless, the crash rate for *LV decelerating* conflicts is zero versus 7 for each of the other conflict types.

One possible reason for this result is based on vehicle kinematics. Since the lead vehicle is still moving in the *LV decelerating* conflicts, the following-vehicle driver has more time to perceive the closing rate and adjust vehicle speed accordingly. A second reason for the result, and one that most likely plays a major role, is that drivers are simply better at perceiving the distance of a moving vehicle versus a stationary vehicle. In the case of the *LV stopped ≤2 s*, drivers do not appear to be particularly adept at judging the point at which the lead vehicle came to a complete stop. Another potential explanation is that drivers are not adept at judging very slow moving lead vehicles just prior to stopping.

With rear-signaling systems, the implication for design is that following drivers need information to help them discriminate between lead vehicles that are moving slowly, moving very slowly (i.e., rolling without use of the accelerator), and stopped. Additional analysis of the RE conflicts should be conducted to examine the relationship between each of these lead-vehicle kinematic conditions and the resulting driver response. Furthermore, the relationship of variables such as traffic density, presence of a junction, and others should be evaluated.

The types of analyses applied in this chapter were also conducted by Tijerina (2004) in his application of the hazard analysis technique to data collected in the ADVANCE study (Dingus, 1997; Dingus et al., 1995). Tijerina's application of the technique was unsuccessful; however, there are a few reasons for this lack of success. First, as noted by Tijerina, the estimation of exposure was weak, since the ADVANCE database contained only 487 vehicle miles of data while the 100-Car Study collected 1.84 million vehicle miles.

Second, the crash data used in the analysis by Tijerina were taken from archival records for the preceding year based on 75.5 million vehicle miles. This is in contrast to the 487 vehicle miles for the near-crash and incident data. For the analyses in this chapter, the event data were taken

from the same database using the same drivers; therefore, the rates from the 100-Car Study are less prone to error. This inference seems to be confirmed through inspection of the confidence intervals for each rate estimate. As opposed to the Tijerina example, no rate estimates for which events were observed have confidence intervals that span zero.

CHAPTER 13: GOAL 9, EVALUATE THE PERFORMANCE OF THE HARDWARE, SENSORS, AND DATA COLLECTION SYSTEM USED IN PHASE II

DATA ANALYSIS OVERVIEW

The 100-Car Study was originally planned as a pilot test to prepare for a Phase IV large-scale naturalistic data collection effort. As such, the purpose of Chapter, 13, *Goal* 9 is to evaluate the hardware, sensors, and data collection systems used in the 100-Car Study and, based on system performance and results, determine the best data collection configuration for a large-scale effort. The definition of what constitutes a large-scale effort has not yet been precisely defined, but, for the purpose of these analyses, the assumption is that 5,000 vehicles will be on the road for a period of two years each.

When reviewing this chapter, it is important to note that the results are very closely tied with the results of the *Goal 10 Report* (separate report) in making suggestions for a future Phase IV effort. This research goal describes hardware, sensor, and data collection system performance throughout the data collection effort in terms of its reliability, maintainability, and sensitivity to event detection. In other words, this research goal focuses on the aspects of data collection that are important to a large-scale study while the *Goal 10 Report* (separate report) focuses on the aspects of data reduction important to a large-scale study.

For Chapter 13, *Goal* 9, different analyses were used to address each of the three questions. To address Question 1, which included the assessment of hardware failure rates, frequencies of equipment failures were tallied based on repair records. Downtime for these failures was also estimated to determine failure rates for different system components.

Question 2, which assessed cost/benefit issues, was addressed by determining the net benefit of valid triggers obtained versus all of the cost factors associated with the triggers (including sensor costs, installation time, etc.) in a large-scale study.

Question 3 focused on determination of the criteria for data triggers in a large-scale effort. The assumption was made that it will not be possible to collect continuous data for 5,000 vehicles as was possible with 100 vehicles. Therefore, a scheme must be developed to trigger data collection for a large-scale effort that minimizes misses and false alarms in the dataset. To address this issue, discriminant analyses and logistic regression were used to determine logical trigger settings that contributed maximum event detection without incurring a large number of invalid events. Within the discriminant analysis approach, various techniques were tried, including stepwise selection of variables and the use of costs for particular classification errors.

Data Included in the Analyses

The analyses described in this chapter used data from vehicle repair logs to determine system and component downtime. In addition, the database of reduced events was used to obtain the efficacy of the various sensors and to determine suitable sensors and their corresponding trigger settings for inclusion in a future large-scale data collection effort.

Question 1. What were the failure rates for each sensor type and/or hardware equipment?

There are several pieces of information needed to address failure rate for each piece of hardware, including:

- When was each vehicle on the road?
- What lessons were learned during the data collection period that improved detection and repair time and when were these lessons learned?
- What kind of failures occurred?
- When did failures occur?
- When were failures detected?
- What were repair times for failures as well as the failure rate for components?

The answers to these questions are not independent. For example, the repair time for each failure was greatly dependent on the lessons learned during the data collection period that improved the efficiency with which repairs were detected and made. The time to make a repair changed depending on when it occurred during the course of the study. In addition, other factors affected the estimation of reliability since, for example, a failure was not detected the moment it occurred. Instead, a failure was detected via three different methods that occurred at different times during data collection. Thus, reliability must be estimated and the questions that were considered to make reasonable reliability estimates are considered in more detail below.

When was Each Vehicle on the Road?

The first vehicle was put on the road during January 2003. From that point on, vehicles were phased into the study until June 2003. At this point, all 100 vehicles were on the road collecting data. In January 2004, the phase-out process began with the vehicles that had been in service for one year. The phase-out process concluded in July 2004, after private vehicles had been on the road for a minimum of one year and leased vehicles had been on the road for a minimum of 13 months. The following table shows the number of vehicles on the road for each month from January 2003 to July 2004.

Table 13.1. Number of vehicles on the road for each month of the study.

Month	Cars on the Road (@ end of month)
January, 2003	6
February, 2003	12
March, 2003	33
April, 2003	65
May, 2003	100
June, 2003	100
July, 2003	98
August, 2003	100
September, 2003	100
October, 2003	99
November, 2003	98
December, 2003	95
January, 2004	96
February, 2004	91
March, 2004	85
April, 2004	67
May, 2004	50
June, 2004	26
July, 2004	0

What Lessons Were Learned During the Data Collection Period That Improved Detection and Repair Time and When Were These Lessons Learned?

A very important factor in determining the severity of a particular hardware failure is the amount of total downtime for the failure. The total downtime is a function of:

1) The time it took to detect the failure;
2) The time to obtain a replacement part(s); and
3) The time to actually perform the repair needed within the vehicle.

Throughout the 100-Car Study, lessons were learned that reduced these times and made data collection more efficient. The lessons learned to reduce the time to obtain a replacement part and to perform repairs are presented in this subsection. Efforts to reduce the time it took to detect a failure are discussed in a different subsection (see the subsequent question in this section called "When were failures detected?").

For the three times affecting a failure described above, the greatest variable tended to be the time required to perform the repair. This timeliness factor was considered an important area in reducing downtime throughout the data collection effort, and was therefore the focus of many timesaving measures. The time to acquire repair parts was also minimized to the extent possible within the study. The lessons learned while developing these timesaving measures were initially discussed in Chapter 4, Lessons Learned, but are expanded here to provide a more detailed framework to support the assumptions made later in the section to calculate estimated failure rates for various components.

A lesson learned from the beginning of the study was that an effective way to minimize repair time was to have replacement parts on-hand at all times. A key lesson learned, however, was the number of spare parts needed in inventory. For example, it was not anticipated that a larger number of radar antennas would be needed due to damage in rear-end crashes. Throughout the study, the time required to acquire a part was eliminated by having a sufficient number of spare parts and/or spare systems available when they were needed. Throughout the data collection effort, there was not a single repair instance in which downtime was increased due to the unavailability of a particular part.

The replacement time was the most variable and lengthy aspect of the failure downtime. For an estimated 50 percent of the duration of the study, the downtime for failures was approximately 2 weeks, since bi-monthly trips were scheduled for technicians to travel from Blacksburg to northern Virginia to fix all failures that had been identified up to that point in time. This downtime period applied to all failures types: catastrophic, major, and minor. During this time period, methods to increase cost-effectiveness, decrease replacement downtime, and reduce the number of active service requests were being explored.

For the remainder of the study, two important measures were taken to reduce the replacement time for inoperative hardware. First, chase vehicle drivers (or data "downloaders") who were already in close contact with the cars were trained to perform minor maintenance procedures after gaining access to the car. Up to this point, the downloaders were only required to notify any perceived faults with the equipment. While effective for fault detection, it was felt that this approach was time-inefficient and under-used downloader abilities. Therefore, for the second half of the study, downloaders were also required to watch video from each vehicle when they downloaded it and fix any cameras that were inoperative. This approach reduced the response time to minor failures from 2 weeks, up until that point, to a maximum of 1 week. Second, an on-site (i.e., a northern Virginia area) technician was hired to respond to any catastrophic or major failures. In addition, the on-site technician was supplied with replacement systems that could be "line-replaced" to decrease overall system downtime. These actions reduced the response time to catastrophic and major failures from two weeks to a maximum of 1 week for the second half of the study.

A logical question is why these lessons took several months to be implemented. Recall that vehicles were phased into the study from January 2003 to June 2003. Until most of the vehicles were on the road, it was possible to make repairs quickly. However, once most of the vehicles were on the road, the full burden of repairs became apparent. It was at this time that new measures to address repairs were adopted.

These lessons learned can be used to estimate system downtime. For approximately the first half of the study, the maximum system downtime (combining fault detection and repair) was less than three weeks. For the final part of the study, the maximum system downtime was reduced to less than two weeks. These downtime distributions will be considered when estimating failure rates for the DAS hardware.

What Kind of Failures Occurred?
Components of the data collection system can be classified into two general categories: data gathering and data storage (in-vehicle). Both of these functions had several distinct hardware items contained within the data acquisition system. For example, each sensor component had three parts: the sensor itself, its associated cable, and the component inside the system box that controlled its data collection (see Chapter 2, Method). The failure of any of these components resulted in a sensor failure. All of these hardware components were subject to failure within the data collection period. However, the failure of a particular hardware component was only as significant as the type of data that was lost due to the failure.

Based on the type of data lost, three different classifications for the observed failure can be considered: catastrophic, major, or minor. Catastrophic failures resulted in the loss of data already stored within the system and the cessation of further data collection. Major failures caused the loss of a substantial number of data streams without loss of data already stored within the system. For example, the system stopped collecting video due to a malfunction in the video board. However, other data (e.g., throttle position, radar) continued to be collected. Finally, minor failures caused the loss of a small number of data streams without loss of data already stored within the system. For example, the face camera view was not acquired, but the remaining camera views were available within the data stream.

There were no catastrophic and major failures requiring a complete system overhaul or replacement of the DAS during the course of the study. Without exception, these failures were repaired through the replacement of one or two parts of the DAS, which was then put into service again. System remove/replace operations to fix these failures occurred in 45 instances. The replacement process entailed the removal of the data acquisition system from the vehicle and installation of a different DAS. The malfunctioning part within the unit was then repaired offline and the refurbished unit was subsequently used as a replacement.

Based on repair logs and communications between technicians, it is estimated that approximately 45 catastrophic or major failures occurred within the data collection period. Causal factors for each of these failures are presented in Table 13.2, along with the sensor or subsystem affected and the frequency of occurrence. Note that the frequencies add to more than 45 because in many instances failures occurred on more than one sensor or subsystem at the same time.

Table 13.2. Causal factors for catastrophic and major failures.

Failure	Instances	Sensor/Subsystem Affected
System fuse blown	5	Power Control Battery Backup
Car battery draining Car failed to start System fuse pulled to prevent battery drainage	28	Power Control Battery Backup
System shuts down during a trip Empty data files	34	Acquisition Software
Hard drive failure Hard drive becomes corrupt	33	Acquisition Software
Download cable failures resulting in incomplete downloads	17	Remote Download
Video problems – No video	22	Real-time Video

Minor failures were, for the most part, constrained to sensors, since loss of data acquisition capability entailed a catastrophic or major failure. A total of 268 minor failures were recorded in repair logs or communications between technicians. These failures, their corresponding sensor or subsystem, and their number of occurrences, are listed in Table 13.3.

Table 13.3. Causal factors for minor failures.

Failure	Instances	Sensor/Subsystem Affected
Improper camera orientation	34	Real-time Video
Cameras falling from mount	21	Real-time Video
Video problems – Single view (other views available)	42	Real-time Video
VORAD – Radar unit	8	Headway Detection
VORAD – Board / Cable	37	Headway Detection
Network box	43	Vehicle Network
Lane tracker	46	Lane Tracker
Cell phone antenna	8	Remote Vehicle Tracking
Other cables	3	---
Internal (backup) battery	6	Power Control Battery Backup
Broken License Plates	20	---

When did Failures Occur?

Extensive logs detailing the timing of fault detection and fault repair were maintained throughout the study. Assuming that fault detection took a set amount of time (i.e., a week), these logs can help establish the timing of the repairs. No particular differences between catastrophic and major or minor failures were observed in terms of timing. Table 13.4 shows the relative number of failures detected during each month of the study. The first failures were detected in March, 2003, and the last failures were detected a year later. The month with the largest percentage of failures was October 2003. Also note that the percentages increased initially as the number of vehicles on the road ramped-up, and decreased as the vehicles were removed from the road after January 2004.

Table 13.4. Percentage of failures for each month of the study.

Month	Percentage of failures
March, 2003	2.3
April, 2003	3.6
May, 2003	6.6
June, 2003	9.1
July, 2003	8.4
August, 2003	7.9
September, 2003	9.1
October, 2003	19.5
November, 2003	7.9
December, 2003	13.7
January, 2004	4.6
February, 2004	3.8
March, 2004	3.6

When were Failures Detected?
The time needed to detect a failure was reduced by requesting this information from various sources. Data downloaders were required to identify, via a checklist, any aspects of the installation that were visually askew as they downloaded data from the vehicle. Furthermore, participants were also instructed throughout the study to call the study contact person when they believed problems existed with any part of the data acquisition system or with the system's interaction with their vehicle. In addition, as data was downloaded and transferred to the storage server, a data reductionist was required to observe a subset of the data and point out any problems noted. It is estimated that these three failure identification methods resulted in a failure occurrence to a failure identification average lag time of one week. Once the problem was identified, total downtime became a function of the availability of replacement parts and the replacement time.

What was the Repair Time for Failures and the Failure Rate for Components?
Out of the 45 estimated catastrophic or major failures outlined above, it is estimated that 20 occurred during the first half of the study (response time less than three weeks) and 25 during the second half (response time less than two weeks). Throughout the study, a total of 4,554 vehicle-weeks were collected from a total of 102 vehicles. Thus, the first half of the study resulted in 60 vehicle-weeks of downtime (20 instances X 3 weeks downtime) while the second half resulted in 50 vehicle-weeks (25 instances X 2 weeks downtime). As a result, the total data collection downtime due to catastrophic or major failures was 110 vehicle-weeks. This represents an overall catastrophic of major failure rate of 2.4 percent (110 vehicle-weeks downtime / 4,554 total vehicle-weeks). The 110 vehicle-weeks of downtime represent, based on an assumed weekly mileage rate for the study of 404.0 miles/vehicle-week (assuming 1.84 million VMT for the study and 4,554 vehicle-weeks), a total of 44,444.4 miles of data that were not collected due to catastrophic or major failures.

Catastrophic and major failure rates per sensor or subsystem were derived from Table 13.2 and are shown in Table 13.5. Sensors and subsystems not mentioned in the table did not exhibit any catastrophic or major failures. A total of 4,554 vehicle-weeks of data collection were used in the calculations. In addition, three weeks of downtime is assumed. This assumption is based on

adding estimates for the time required to detect a failure (~1 week) and estimates for the time to perform a repair (~2 weeks). This estimate is somewhat conservative, since in many instances it took fewer than 3 weeks to detect and repair a fault, especially in the latter part of the study. Thus, the failure rates presented in this section represent a ceiling for the hardware used in the study.

Table 13.5. Catastrophic or major failure rates by sensor or subsystem.

Failing Sensor/Subsystem	Instances	Failure Rate (%)
Power Control Battery Backup	33	2.2
Acquisition Software	67	4.4
Remote Download	17	1.1
Real-time Video	22	1.4

Minor failure rates per sensor or subsystem, compiled from Table 13.3, are shown in Table 13.6. An assumption of three weeks downtime is used, along with a total data collection period of 4,554 vehicle-weeks. These 268 minor failures represent 804 vehicle-weeks of incomplete data. This means the overall minor failure rate (assuming independent failures and the downtime assumptions used before) was 17.7 percent. A total of 324,816.0 miles of data were incomplete, based on the assumed weekly mileage rate for the study of 404.0 miles/vehicle-week. In some cases, this data could still be used in data reduction because a redundant source of data was available.

Table 13.6. Minor failure rates by sensor or subsystem.

Failing Sensor/Subsystem	Instances	Failure Rate (%)
Power Control Battery Backup	6	0.4
Real-time Video	97	6.4
Headway Detection	45	3.0
Vehicle Network	43	2.8
Lane Tracker	46	3.0
Remote Vehicle Tracking	8	0.5

Overall, all failure rates were relatively low. None of the component specific rates were larger than 10 percent and the majority of the rates were smaller than 5 percent. In addition, many subsystems (i.e., accelerometer, critical incident button, gyroscope, GPS) had overall failure rates of zero.

Assuming that all failures resulted in loss of data, the maximum number of miles lost due to failures would have been near 370,000 miles (the sum of miles lost due to all types of failure). There are data for 1.37 million miles, with an estimated total number of miles of possible data available equal to 1.80 million, a difference of approximately 430,000 miles. The difference between the miles lost due to failure (370,000) and the estimated total miles lost (430,000) is 60,000 miles. This discrepancy is probably due to log discrepancies and a few outlier cases in which the failure detection took an inordinately long time for one reason or another. Overall, however, this comparison justifies the validity of the rates calculated in this section and serves as an error check for any repairs that might have been missed in the logs. While it is possible that

some repairs were never entered, the agreement between miles calculated using different approaches supports that the number of these missing repairs would be relatively small (i.e., 60,000/1,800,000 x 100 = 3.3%).

It is important to note that minor failures were also assumed to result in loss of data. However, certain types of failures, such as vehicle network or network box problems, resulted in data that were not analyzed because the trigger criteria relied upon the data generated (e.g., speed). This means that a significant portion of this data can potentially be recovered in the future by estimating the associated parameters *post hoc*.

The failure rates discussed in this section can now be combined with the benefits that would be expected from each of the subsystems within a large-scale naturalistic study. These benefits are discussed in Questions 2 and 3 from two different perspectives. Question 2 discusses the benefits that could be expected from different sensors, hardware components, and data collection components when compared to their failure rates. Question 3 addresses how these systems can be optimized so that the least amount of hardware produces the best possible results.

Question 2. What is the relative cost/benefit for each sensor type and/or hardware component?

A large naturalistic study with 5,000 cars on the road for two years would raise a series of unique and important issues. The larger scale of the project makes many of the support mechanisms available during the 100-Car Study not feasible due to: (1) the large geographical area in which a study like this would need to be conducted, and (2) the sheer number of vehicles and drivers that would have to be tracked.

Perhaps the most important of these issues is that chase cars with downloaders would not be feasible. Maintaining a fleet of chase cars and hiring the personnel required would be too costly and time-consuming. Thus, it is foreseen that cars would be released on the road for a period of six months. At the end of six months of data collection, vehicles would return to have the data collection system removed. At that time, the data stored on the vehicle during the data collection period would be downloaded and backed up. Note that there would be an exception to this scheme when a crash was detected. It is envisioned that automated crash detection would be available as part of a 5,000-car system and when a crash was detected, the data would be immediately downloaded.

This approach poses several challenges. First, the DAS within the car must be highly reliable and resistant to data corruption. Otherwise, it would be highly possible that data from many vehicles would be lost due to system failure. Second, the sensors and hardware associated with the system must have low failure rates to increase the opportunity of acquiring the most data possible during the study period. If any of the sensors is failure-prone, then it will be more likely to fail during the course of the study. Since there would be no interaction with the DAS for six-month intervals, sensor failures would not be detected and any events occurring after the sensor failure would be lost permanently. Third, the number of sensors and other hardware components should be reduced to the largest extent possible in order to minimize failure rates and in-vehicle

data storage needs. A greater number of sensors would increase the likelihood that the DAS would run out of in-vehicle storage space before the end of the data collection period.

In-vehicle data storage capacity needs can also be addressed by considering the amount of data that are actually collected. Since it was considered to be a pilot study, the 100-Car Study used continuous data collection. In this case, continuous collection was needed to observe all driving behaviors and actions in order to determine which of those actions merited more detailed consideration from a traffic safety perspective. For a large-scale study, this would not be feasible or needed. Thus, data collection would only occur with specific triggers. The development of these triggers and their effectiveness is the focus of Question 3 of this goal and discussion of this issue will be deferred until then. The effectiveness of the sensors and hardware suggested for use on a large-scale study at the end of this section have been verified by the results presented in Question 3.

The number of sensors and other hardware should also be reduced to decrease system failure rates and installation requirements. It is logical to expect that the fewer sensors and other hardware components that are needed, the fewer opportunities there will be for the system to break down and stop collecting data. In addition, fewer hardware components also imply fewer wires and fewer control boards within the DAS, which implies easier installation and perhaps even a smaller DAS form factor. Fewer hardware components also mean that installation might be possible on a larger variety of vehicle makes and models, an essential aspect of a 5,000-vehicle study.

DAS components can be split into three categories for this discussion: in-vehicle data storage, sensing, and other hardware (e.g., video). The benefits for each of these components have to be considered separately. In some cases, the benefits can only be presented subjectively, while in other cases, quantification of the benefits is possible.

In-vehicle data storage was an essential part of the 100-Car Study and would also be a central aspect of any large-scale study. The goal of any of these studies is to obtain data from drivers in a naturalistic setting. If these data are lost, the study has been unsuccessful. Thus, the benefits of data storage are quite large, albeit difficult to quantify, and efforts should be directed toward improving the reliability of this system to the degree possible. For the 100-Car Study, catastrophic and major failures occurred at a rate of 2.4 percent. For a 5,000 vehicle study, this would imply that 120 cars would lose their ability to collect or store data throughout the study. Thus, improvement in this figure is probably desirable.

The next important component of any large-scale DAS is the array of sensors included within the system. The benefit for each of these sensors relies on the number of triggers that the sensor is able to support, as well as the effectiveness of these triggers. Maximizing these aspects of the DAS is the focus of Question 3; however, these aspects are initially discussed here because they are directly relevant to determining the benefit provided by each of the sensors that could be used on a large-scale naturalistic study.

Each of the events (i.e., crashes, near-crashes, and incidents) that were collected for the 100-Car Study was selected as a function of one or more triggers. These triggers, in turn, depended on

the sensors or subsystems that collected the data against which the triggers were contrasted. This section examines the effectiveness of each trigger type (and thus each underlying sensor or subsystem type) in correctly identifying a valid event. More details about each of the triggers can be found in the discussion for Question 3.

The database of reduced events included 69 crashes, 761 near-crashes, and 8,295 incidents. Table 13.7 lists each of the sensors or subsystems used in the data acquisition system and the percentage, by severity, of the valid events detected by each trigger. These parameters are indicative of the first measure of benefit for a trigger, which is the proportion of valid events that would be detected. In a large-scale study without continuous data collection, a failure of a trigger to activate for an event would mean the loss of the event. Thus, the higher the percentage of valid events that a trigger captures, the better that particular trigger is considered to be. Note that each sensor is independent in the number of events that are detected, meaning that the percentages are not additive for any column.

Table 13.7. Sensor or subsystem benefits in terms of valid events captured.

	Severity			
Sensor/Subsystem	**Crashes Detected** *% (Instances)*	**Near-crashes Detected** *% (Instances)*	**Incidents Detected** *% (Instances)*	**Valid Events Detected*** *% (Severe Instances) (All instances)*
Accelerometer (Lateral)	18.8% (13)	3.2% (24)	0.6% (48)	3.5% (85) (316)
Accelerometer (Long.)	58.0% (40)	61.9% (471)	37.6% (3,089)	44.7% (3,600) (4,078)
Critical Incident Button	18.8% (13)	16.0% (122)	5.2% (434)	8.4% (569) (762)
Range/Range Rate Detection – Fwd.	24.6% (17)	42.8% (326)	57.5% (4,768)	56.5% (5,111) (5,158)
Range/Range Rate Detection – Rear	2.9% (2)	7.6% (58)	4.4% (363)	4.6% (423) (424)
Gyroscope (Yaw Rate)	24.6% (17)	25.0% (190)	17.2% (1,423)	21.7% (1,630) (1,983)
Lane Tracker	0.0% (0)	0.5% (4)	0.1% (5)	0.6% (9) (82)
Radar – Side	0.0% (0)	0.3% (2)	3.0% (251)	3.1% (253) (280)

*Severe instances include crashes, near-crashes, and incidents. All instances also include non-conflict incidents. The percentage provided is based on all instances.

Only the forward headway detection sensor was able to identify more than 50 percent of the valid events (56.5%), with the longitudinal acceleration sensor identifying the second largest percentage of valid events (44.7%). When event severities are considered, some could be predicted more than 50 percent of the time by the data from some sensors. The longitudinal accelerometer sensor performed best in crashes, correctly marking as an event 58.0 percent of all of the crashes in the dataset. This sensor was also the most effective in identifying valid near-crash events (61.9%). The forward headway detection sensor correctly triggered for 57.5 percent of all incidents. Finally, all three severity levels had reasonably high detection rates triggered from the gyroscope sensor.

It is important to note that the lane tracker and side radars are not fairly represented in Table 13.7. The side radars were only present on one fifth of the fleet (20 leased vehicles) for 6 months. Thus, their rate should be less than 10 percent of any of the other sensors. However, the inclusion of side radars for a full-scale 5,000 vehicle fleet would probably be impractical due to cost and installation requirements.

However, the lane tracker may present a different case. The lane tracker is software-based and uses the same forward camera already present for the study. While it does increase computer processing requirements, the cost is fairly minimal. In addition, due to the focus of the current study on rear-end crashes, there was not a great deal of effort made to determine the feasibility of using the lane tracker as a trigger. Early attempts did show that the lane tracker signal was noisy both for reasons of road marking visibility and driver behavior (i.e., many drivers exceeded lane boundaries on purpose or relaxed standards in the absence of other traffic). Therefore, a lane tracker may turn out to be a very valuable addition to a study that is more broadly focused and when more time and resources are available to improve the signal filtering.

Another measure of benefit from a trigger relates to its ability to capture only valid events, which in turn indicates the level of noise (i.e., invalid events) that will exist within the stored data in a large-scale study. Triggers that perform poorly in this regard will overload the in-vehicle data storage equipment with useless data and could compromise the capacity of the system to store important events occurring after the in-vehicle data storage unit is full. Table 13.8 shows the rate at which invalid events were found for each trigger. A lower rate indicates better trigger performance for this measure. In addition, the catastrophic/major and minor failure rates were noted. Note that each sensor is independent regarding the number of events detected; thus, the percentages are not additive for any column. Also note that for the side radar, the same failure rates used for headway detection are used, given that the same technology was applied.

Table 13.8. Sensor or subsystem benefits in terms of proportion of invalid events.

Sensor/Subsystem	Invalid Events Detected % *(Instances)*	Failure Rate (%) (Catastrophic & Major / Minor)
Accelerometer (Lateral)	91.3% (3,325)	0 / 0
Accelerometer (Long.)	66.4% (8,047)	0 / 0
Critical Incident Button	69.9% (1,773)	0 / 0
Headway Detection – Fwd.	83.4% (25,833)	0 / 3.0
Headway Detection – Rear	59.9% (633)	0 / 3.0
Gyroscope (Yaw Rate)	91.1% (20,217)	0 / 0
Lane Tracker	96.1% (2,532)	0 / 3.0
Radar – Side	96.5% (13,808)	0 / 3.0

These higher sensitivities to valid events were not necessarily accompanied by good negative predictive ability. The forward headway detection sensor had a rate of invalid events of 83.4 percent, near the top of the list for this category. The longitudinal acceleration sensor fared better, with 69.9 percent of all the events that it triggered being classified invalid (second from the bottom of the list for this category). The rear headway detection sensor had the lowest rate of invalid events, at 59.9 percent.

Perhaps the single most important sensor present in the 100-Car Study data collection system was video. The benefit and value of video is very large, given the aid it provides to data reductionists in determining the validity of triggered events. Given that invalid events seem unavoidable based on current sensor technology (see the discussion for Question 3 for more on this issue), it seems that some video views would be absolutely necessary in a large-scale naturalistic study to aid in the data reduction process. This issue is discussed in more detail under the *Goal 10 Report* (separate report). However, note that even with a failure rate of 6.4 percent, video is considered an important piece of hardware to be included on a large-scale study due to its data validation benefit. A logical concern would be that with this relatively high failure rate, video might become problematic to include in a large-scale study. While this concern is justified, the failure rate can be reduced by including only camera views that are absolutely necessary (which may also reduce in-vehicle data storage requirements). Justifications for the elimination of certain camera views are presented in the *Goal 10 Report* .

Other subsystems not shown in Tables 13.7 or 13.8 were not used to collect any trigger information or validate events. Rather, they served to support the data collection function or the remote vehicle tracking function. These functions contributed to the success of the overall system and their absence in this analysis simply indicates that they had no direct bearing on the selection of particular events for further analysis, which is the main source of benefit information. For example, RF and glare sensors were installed in vehicles, but their data were

not considered essential or useful in triggering for events. These sensors are part of a category that can be deemed optional, whose use depends on the specificity of the data desired within any larger-scale study. Their usefulness and corresponding benefit lies in providing the ability to characterize events better than by simply looking at a video. However, this ability has to be weighed against their cost, expected effectiveness, and required maintenance. Whereas the effectiveness for the sensors considered above in this section can be determined based on their performance, the effectiveness of optional sensors is more subjective, and depends on the value of the information to the stakeholders for a particular data collection effort. Costs of maintenance did not seem to be high, based on their absence from repair logs. However, some of these sensors can be noisy, and care should be taken in ensuring that the quality of data from any of them included in a large-scale study is sufficient to justify their expense in terms of cost and of in-vehicle data storage.

The data available for all of these in-vehicle data storage components, sensors, and hardware components are limited in terms of benefits, since most of the benefit gathered from the components is difficult to quantify. The costs, in terms of failure rate (i.e., required maintenance and repairs), in-vehicle data storage needs, and classification effectiveness, while somewhat more quantifiable, are also subject to some degree of interpretation.

When benefits and costs are observed as a whole, several technologies stand out for inclusion in a large-scale naturalistic study. Accelerometers, yaw rate sensors, and range/range rate sensors (particularly forward) seem essential to the real-time classification of valid events for a larger-scale study using a trigger-based data collection system. Video is also necessary to allow for screening the invalid events that will inevitably be collected. If these technologies are combined using algorithms that aggregate their data (as discussed in Question 3), they should be able to collect an acceptable number of valid events while minimizing the number of invalid events that are stored and have to be eliminated manually.

Question 3. Based on data collection in Phase II, what are the optimum sensor values that should be used for a triggered data collection system?

As discussed in Question 2, an important aspect of a large-scale naturalistic data collection effort would be its trigger-based nature. While the 100-Car Study collected data continuously, a larger-scale study would not likely have the storage and data management support structure that the 100-Car Study had. Question 2 already evaluated the sensors and equipment needed to create these triggers in terms of their individual contributions to the event detection task. More information on that process is presented here and augmented with techniques to aggregate the data from various sensors in order to further improve event classification accuracy.

Thus, the goal of this sensitivity analysis was to determine an array of sensor and/or sensor thresholds that could be used to identify the largest possible number of crashes and near-crashes while minimizing the number of invalid events that are misclassified. As has been discussed elsewhere, equipment within the car was able to measure lateral and longitudinal acceleration, yaw rate, forward time-to-collision, rear time-to-collision, headway, and lane busts. Each of these measures was foreseen as a possible indicator of a crash or near-crash under particular

circumstances, but the appropriateness of this sensor suite and potential threshold settings had to be determined using the empirical data available.

This section also describes, when feasible, trigger performance as a function of crashes, near-crashes, and incidents. Note that a large number of incidents were recorded and compared to the number of crashes and near-crashes obtained in the study. It is possible that incidents would not be desired in a large-scale study in order to reduce the total number of events (valid and invalid) for which data are collected. This issue is also explored within this question.

Trigger-Based Sensitivity Analysis

Initial work to address this question leveraged on the activities performed to define the set of trigger settings used to select the events that were reduced by analysts. These activities involved the creation and testing, within a limited initial dataset, of several trigger settings. The goal of these initial analyses was to minimize the number of missed events with less regard for the number of invalid events that were selected. These analyses provided substantial insight into the arrangement of the data and had important implications for the methods used to carry out the sensitivity analysis for the larger subset of the data.

The results of these analyses were used to synthesize a set of triggers that are described in Table 13.9. These triggers were used for the detection of events within the full dataset. Each table entry also presents, broadly, the characteristics of other less effective iterations of the triggers, based on the initial testing performed within a limited dataset. Unless otherwise noted, triggers were only recorded when the vehicle was moving (speed greater than 0.0 m/s). The effectiveness obtained from each of these triggers is repeated here and it was also addressed in Question 2. These effectiveness values, initially discussed in Question 2, represent reasonable estimates of the maximum effectiveness that can be obtained with each trigger. They also serve as an upper limit on the effectiveness that can be obtained from considering sensor outputs independently.

Unfortunately, the effectiveness numbers for each trigger are substantially lower than the 90 percent valid event retention and 5 percent invalid event retention that were originally proposed as goals for a large-scale study. Fortunately, as will be discussed in the next section on data reduction, a larger percentage of invalid triggers add only marginal cost to a large-scale effort. The downfall of simple, highly effective trigger criteria is the amount of individual variability present in light vehicle driving. (Prior studies, such as Dingus et al., 2002, had greater success with heavy vehicles.) This variability proved to be too large to be handled by set triggers based on a relatively simple logic. Statistical methods to improve these classification accuracies by using aggregate sensor data are discussed in the next section.

Table 13.9. Trigger characteristics for the 100-Car Study.

Trigger Type	Description	Unsuccessful Alternative Trigger Setting Tried	Valid Events Detected (%)	Rate of Invalid Events within the Trigger (%)
1. Lateral Acceleration	• Trigger setting had to be set very high to avoid the misclassification of a large number of invalid events. • The final trigger criterion was set at ± 0.7 g • Suitability of this trigger might improve with the presence of side sensors that aid in identifying adjacent vehicles in lane change or merge events (see Side Sensor Triggers section below).	• Data filtering • Smaller threshold values	3.5	91.3
2. Longitudinal Acceleration	• Vehicle's longitudinal acceleration reached a value ≤-0.52 g at any single point in time, with other levels of acceleration contributing to a trigger depending on data from other sensors (see the discussion for the Forward TTC).	• Data filtering • Less extreme thresholds (e.g., -0.3 g) • Using acceleration measured over a larger sample (e.g., 5 samples)	44.7	66.4
3. Event Button	• This trigger occurred whenever participants pushed a button within the cabin. The trigger depended on driver willingness to inform the researchers about the occurrence of an incident, and could be noisy due to drivers abusing of the feature to point out irrelevant events.	• N/A	8.4	69.9
4. Forward Time-to-Collision	• Forward TTC value was based on the standard TTC equation using range, range rate, and acceleration • Filtered to exclude: o Any approaching object. o Any triggers that had a corresponding yaw signal that was >4 degrees per seconds and	• Speed dependent TTC boundary line.	56.5	86.4

	occurred simultaneously (within 3 s) with the Forward TTC trigger while the vehicle was traveling at 8.9 m/s (20 mph) or less (i.e., indicating vehicle turning). o Any signal from the radar that did not maintain a consistent signature for a minimum of 7 frames. o If an initial positive Forward TTC trigger was observed, the peak longitudinal acceleration was determined across a 6 seconds time sample, 3 seconds prior to and following the Forward TTC trigger. o All peak longitudinal accelerations smaller than -0.5 g that were coupled with a Forward TTC of 4 seconds or less represented a trigger. o All longitudinal accelerations between -0.4 g and -0.5 g represented a trigger provided that the Forward TTC value was ≤4 seconds and that the corresponding forward range value at the minimum Forward TTC was ≤ 30.5 m (100 ft).			
5. Rear Time-to-Collision	• Calculated using range, range rate, and acceleration • Ignored targets with a speed ≥44.7 m/s (100 mph) • Used a trigger value of two seconds or less, as long as the corresponding rear range was ≤15.2 m (50 ft) and the peak longitudinal acceleration of the following vehicle was <-0.4 g. The peak longitudinal acceleration of the following vehicle was sampled at +/- 2 seconds around the TTC trigger.	• Considering Rear TTC values by themselves, without consideration of the following vehicle deceleration or speed.	4.6	59.9
6. Yaw rate	• Used to identify a change in heading	• Speed dependent boundary line.	21.7	91.1

	that was immediately followed by a return to the same general heading. • This filter was suitable for determining when the driver performed a sudden steering maneuver and not a cornering maneuver. • The trigger criterion for yaw rate was any set of values that went from neutral (i.e., ~0) yaw rate to +4 degrees/s, oscillated back to -4 degrees/s (or vice versa: -4 to +4), and then returned to neutral within a 3-second time window. Thus, the vehicle was required to return to the same general direction of travel within the 3-second window. • A minimum speed of 6.7 m/s (15 mph) was required for the trigger to activate.			

The triggers in Table 13.9 represent the best performing triggers, but there were other triggers that resulted in detection of very large numbers of invalid events, or for which only a limited amount of data were collected. These triggers included criteria from the lane tracker and the side radars.

The lane tracker resulted in valuable data used during the data reduction process to understand both the type and severity of the events. In addition, triggers were attempted to capture lane abort and lane bust events. A lane bust occurred when the vehicle crossed a solid lane line. This trigger was set to occur when the vehicle moved a minimum of 3 ft outside of a lane boundary while traveling at a speed ≥20.1 m/s (45 mph) and underwent several unsuccessful revisions. The lane abort trigger was activated when a vehicle crossed a dashed line and returned to the original position. This trigger also underwent several revisions that failed to improve its performance. The lane abort trigger was set to occur when the vehicle moved a minimum of 3 ft outside of a lane boundary (60 in from center of lane) while traveling at a speed ≥20.1 m/s (45 mph) and the vehicle did not complete the lane change. Altogether, these triggers correctly classified 0.6 percent of all valid events. A total of 96.1 percent of all events captured by this trigger were invalid.

The vehicles that were instrumented with side sensors allowed the establishment of additional triggers that provided more detailed information on side conflicts. Four separate triggers were developed based on the data for these sensors:

- Turn Signal Light: This trigger occurred when an object was detected by the side radar within +/- 1 second of any instance in which the turn signal was active. Vehicle speed also had to be higher than 6.7 m/s (15 mph) for the trigger to occur.
- Side Cutoff: This trigger was only activated when vehicle speed was faster than 8.9 m/s (20 mph) and a lane change occurred in front of another car located within 15.2 m (50 ft) of the vehicle.
- Side Blind Spot: This trigger occurred only for vehicle speeds >8.9 m/s (20 mph) when a lane abort maneuver (as detected by the lane tracker) occurred while an object was detected by the side radar.
- Side Yaw: The trigger criterion for yaw rate was any set of values that went from neutral (i.e., ~0) yaw rate to +2 degrees/s, oscillated back to -2 degrees/s (or vice versa: -2 to +2), and then returned to neutral within a 3-second time window. A minimum speed of 6.7 m/s (15 mph) was required for the trigger to activate. In addition, the side radar had to detect an object for the trigger to occur.

Altogether, these triggers correctly classified 3.1 percent of all valid events. A total of 96.5 percent of all events captured by this trigger were invalid.

Multivariate Statistical Sensitivity Analysis

Given that the univariate (albeit multiple, sequentially filtered) triggers did not yield the desired levels of identification performance, multivariate statistical approaches were tried. These approaches exploited the relationship between the variables of interest to make classification decisions. These decisions are not dependent on a single pass/fail value for a particular sensor (or combination thereof), but consider the overall contribution of each sensor available in determining the likelihood of identifying a crash or near-crash. In simple terms, these

approaches build complex classification matrices by mathematically synthesizing the data that are available. This complexity also hinders their ease of understanding. While it is easy to imagine that there is a high probability that an event occurred if a longitudinal acceleration greater than 0.9 g is observed, it is less intuitive to understand that a longitudinal acceleration of 0.3 g might indicate an event if the forward time to collision is less than 0.7 seconds, the speed is less than 15.6 m/s (35 mph), and the lateral acceleration is lower than 0.2 g. These complex if-then relationships are synthesized by these statistical approaches and put into the forms of simple linear equations producing the desired classification results.

All of the crashes and near-crashes, as well as a portion of the triggered events that were found to be invalid, were selected for use in these multivariate statistical analyses. Two approaches were considered, based on their adequacy for the problem of interest: logistic regression and discriminant analysis. Both of these approaches are able to take a large number of variables and use them to either classify observations (discriminant analysis) or determine the probability of group membership for any observation (logistic regression). The main difference between these two techniques is that discriminant analysis provides classification for continuous dependent variables that preferably fit a multivariate normal distribution. Logistic regression can use continuous or discrete variables and makes no multivariate normality assumption.

Initial analysis approaches considered the full dataset of valid events, without filtering, by event severity. A random sample of invalid events was also included within the analysis dataset. The sample consisted of 17,625 invalid triggers, representing 12.9 percent of the overall number of invalid triggers in the study.

The dataset included 8 seconds of data per event (valid or invalid), centered around the instant in which the first trigger for the event occurred. Thus 4s of the data for each event occurred before the trigger and 4s occurred after the trigger. If a valid event had more than one positive trigger, the data for the event was centered on the first trigger in the time sequence.

From the 8 seconds of data per event, a series of dependent variables were calculated for inclusion in subsequent analyses. Unless otherwise noted, all of these variables were used in each of the analyses. These variables included:

- Forward TTC considering lead-vehicle acceleration – minimum through the 8 seconds of data for the event.
- Forward TTC without considering lead-vehicle acceleration – minimum through the 8 seconds of data for the event.
- Range at the minimum Forward TTC (not considering lead-vehicle acceleration).
- Rear TTC considering following-vehicle acceleration – minimum through the 8 seconds of data for the event.
- Rear TTC without considering following-vehicle acceleration – minimum through the 8 seconds of data for the event.
- Range at the minimum Rear TTC (not considering following-vehicle acceleration).
- Maximum difference between yaw rates – calculated between the maximum and minimum values obtained for the 8 seconds of data.
- Time between maximum and minimum yaw rates – for the 8 seconds of data for the event.

- Yaw variance – for the 8 seconds of data for the event.
- Longitudinal acceleration – mean, minimum (recall that a deceleration is negative), and variance for the 8 seconds of data.
- Absolute lateral acceleration – mean, maximum, and variance for the eight seconds of data.
- Speed – mean and variance for the 8 seconds of data.

Discriminant Analysis. Discriminant analysis is descriptive of a particular statistical process with a large number of associated options. The discriminant analyses discussed are the end product of a large number of trial runs exploring the options that could best be used to model the data at hand. For example, an important parameter for discriminant analysis is the rule used. The analyses in this section employed a quadratic discriminant rule. For many of the analyses discussed, other rule options, including a linear discriminant rule and nearest neighbor rule (which does not depend on multivariate normality) were tried. No cases were found in which their performance was better than the quadratic discriminant rule.

Since the lack of multivariate normality in the data was a source of concern when applying this procedure, transformations of the variables were also attempted to attain multivariate normality. These transformations did not result in any improvements to the classification accuracy of the discriminant analysis.

Another discriminant analysis option is the method used to estimate the probabilities of misclassification. The analyses in this section employed the cross-validation method because this approach produced nearly unbiased estimates of the true probabilities of correct and incorrect classifications.

Finally, discriminant analysis also allowed the introduction of expected probabilities of events within the dataset. These probabilities, as will be shown later in this section, could have a large effects on the classification accuracies of the procedure. Unless otherwise noted, the analyses discussed assume no knowledge of prior probability, and allow the analysis to use the probabilities that it detected within the dataset as the expected classifier probability.

Some manipulation of the available data also occurred. Since the variables being used were kinematic in nature, it was considered necessary to initially split the data between compact and mid-size automobiles ("cars") and sport-utility vehicles ("SUVs"). These two vehicle types can have differing kinematic behaviors and differing driver responses to control them. Thus, this additional source of variability within the data was isolated by the data split.

The initial discriminant analyses, considering cars and SUVs separately, included all variables except yaw variance, longitudinal acceleration variance, and lateral acceleration variance. These variables were added to later analyses to attempt, unsuccessfully, to improve classification accuracy. For this set of analyses, events of all severities were included as valid events.

Results here and throughout this section are shown in terms of a confusion matrix. The rows in the matrix indicate the correct classification for the event. The columns in the matrix indicate the classification provided by the discriminant analysis. The numbers within the matrix cells

indicate the percentage of observations for a row that were classified in each particular category. The percentages total 100 percent for each row. The table's diagonal (starting at the upper left corner) is bold-faced to indicate correct classifications. Results for the discriminant analysis on data for passenger cars are shown in Table 13.10.

Table 13.10. Confusion matrix for passenger cars.

		Event classified as:	
		Invalid	**Valid**
Event was:	**Invalid**	**76.4**	23.6
	Valid	54.6	**45.4**

Results for the discriminant analysis on the data for SUVs are shown in Table 13.11.

Table 13.11. Trigger Confusion matrix for SUVs

		Event classified as:	
		Invalid	**Valid**
Event was:	**Invalid**	**81.3**	18.7
	Valid	41.2	**58.8**

These numbers were discouraging, as less than 60 percent of valid events could be detected. Perhaps even more discouraging, even at this high miss rate, was that more than 15 percent of all invalid events would have to be reduced, because the system would collect them as valid data. Given the overwhelming presence of invalid events in the data (i.e., in which at least one of the triggers was active in a situation when no event occurred), this would result in a large number of invalid events that would have to be reduced for each valid event (approximately 15:1). However, given the broadness of the data, it was thought that variability due to the nature of the event could be influencing the results. The kinematic conditions for a conflict with a lead vehicle that was a near-crash, for example, could be different from the kinematic conditions for a conflict with a following vehicle that was a near-crash. Forcing the discriminant analysis to lump these two categories into the same general classification (i.e., valid event) could hinder the overall classification ability.

Initial support for this hypothesis was obtained from graphs of the dependent variables for crashes, with crashes grouped by nature. Some differences in vehicle kinematics could be observed as a function of event nature. To determine whether these observations could translate into a better classification scheme, a discriminant analysis was created to classify events based on their nature. Five different nature categories were collapsed from the broader set of categories employed by data reductionists:

1. Conflict with a lead vehicle (CLV).
2. Conflict with a following vehicle (CFV).
3. Side conflict (CS).
4. Conflict with non-vehicle or parked vehicle (CNV).
5. Single vehicle conflict (SVC).

In addition, a sixth category was included that grouped invalid events (INV). This analysis only considered crashes and near-crashes. Any separation in kinematics that might be useful for

classification would be intensified on these event severities (when compared to incidents, the remaining type of valid event). If the analysis showed good classification performance for crashes and near-crashes, it could be extended to include incidents. This philosophy was maintained for the remainder of the analyses in this section.

This analysis also considered both cars and SUVs through the use of an artificial variable within the set of predictors. Thus, possible variance due to vehicle type was accounted for without the need for separate discriminant analyses for each vehicle type.

As for the previous analysis, yaw variance, longitudinal acceleration variance, and lateral acceleration variance were not included. These variables were added on later analyses to attempt, unsuccessfully, to improve classification accuracy. To simplify the matrix, counts are not provided. Table 13.12 shows the results for this discriminant analysis.

Table 13.12. Confusion matrix for different event types.

		Event classified as:					
		CFV	**INV**	**CLV**	**CNV**	**CS**	**SVC**
Event was:	**CFV**	**1.3**	76.6	10.4	0.0	7.8	3.9
	INV	1.0	**92.7**	2.8	0.3	1.8	1.5
	CLV	1.6	82.3	**10.3**	1.6	3.4	0.8
	CNV	2.7	75.7	13.5	**2.7**	0.0	5.4
	CS	2.0	72.6	5.6	1.0	**15.7**	2.9
	SVC	1.5	61.8	0.0	2.9	10.3	**23.5**

While the approach classified a large number of invalid events correctly, it also tended to classify the majority of all other types of events as invalid. Thus, while the number of invalid events that would be observed by data reductionists would be smaller, a large number of valid events would be missed.

At this point, it was possible that prediction of the discriminant analysis could perform better for crashes than near-crashes, so these two severities were separated. In addition, the yaw variance, longitudinal acceleration variance, and lateral acceleration variables were added to the analysis to attempt improving the classification accuracy. As for the previous analysis, cars and SUVs were both included in the analysis via an artificial variable. For this table, the total number of crashes expected was 69 and the total number of near-crashes expected was 761. Results for this discriminant analysis are shown in Table 13.13.

Table 13.13. Confusion matrix for different event severities.

		Event classified as:		
		Crash	**Invalid**	**Near Crash**
Event was:	**Crash**	**15.0**	60.0	25.0
	Invalid	2.5	**94.3**	3.2
	Near-crashes	4.6	64.8	**30.6**

Thus, the discriminant function classified near-crashes two times more accurately than crashes. However, both crashes and near-crashes were classified as invalid close to 60 percent of the time. The difference in classification accuracy between crashes and near-crashes was due to crashes being classified as near-crashes.

For some discriminant analyses, stepwise selection of variables can help by eliminating variables from the analysis that are either not contributing to the discrimination, or that are contributing the same information as other variables. At this stage of the analysis, a stepwise discriminant analysis was performed to determine if this was a possibility with the dataset. The stepwise discriminant analysis retained all of the original variables.

The conclusion from these efforts was that it might be useful to separate and analyze events according to their event nature. These analyses were performed only for conflict with lead vehicle and conflict with following vehicle, as these were the two most frequent categories of valid events. Remaining categories occurred less frequently by a factor of at least two. Again, SUVs and cars were considered within the same analysis via an artificial variable. Results of the discriminant analysis for conflict with lead vehicle are shown on Table 13.14.

Table 13.14. Confusion matrix for conflict with lead-vehicle events.

		Event classified as:	
		Invalid	**Valid**
Event was:	**Invalid**	**89.7**	10.3
	Valid	44.3	**55.7**

Results for conflict with following vehicle are shown on Table 13.15.

Table 13.15. Confusion matrix for conflict with following-vehicle events.

		Event classified as:	
		Invalid	**Valid**
Event was:	**Invalid**	**99.2**	0.8
	Valid	94.9	**5.1**

While the percentages of correct invalid classifications are large, the percentages of incorrect invalid classifications are small, at 55.7 percent for conflicts with lead vehicles and 5.1 percent for conflicts with following vehicle.

A final option to improve the performance of the discriminant analysis was to modify the expected probabilities of valid and invalid events within the dataset. Until this stage, those probabilities were being calculated from the data used. Given that some knowledge of these probabilities was possible, it was foreseeable that they could be set instead of calculated. The analyses that follow consider only conflicts with lead vehicles, as they are exploratory in nature. Similar levels of performance could be obtained for conflicts with following vehicles, if needed, by using slightly different expected probabilities.

These analyses include SUVs and cars via the inclusion of an artificial variable within the analysis. Results of the discriminant analysis for conflict with lead vehicle with equal probability of valid and invalid events are shown on Table 13.16.

Table 13.16. Confusion matrix for conflict with lead-vehicle events, equal event probability.

		Event classified as:	
		Invalid	**Valid**
Event was:	**Invalid**	**41.1**	58.9
	Valid	11.1	**88.9**

In this case, 76.9 percent and 89.3 percent of crashes and near-crashes are classified correctly. If the expected probability of an invalid event is increased to 0.7 (from 0.5 in the previous analysis), the results are shown on Table 13.17.

Table 13.17. Confusion matrix for conflict with lead-vehicle events, 0.7 invalid event probability.

		Event classified as:	
		Invalid	**Valid**
Event was:	**Invalid**	**50.8**	49.2
	Valid	15.1	**84.9**

Now, 76.9 percent and 85.2 percent of crashes and near-crashes are classified correctly. Note that the adjustment in cost did not affect the classification accuracy of crashes. The percentage of correctly classified invalid events increased. If the expected probability of an invalid event is further increased to 0.9 (from 0.7 in the previous analysis), the results are shown on Table 13.18.

Table 13.18. Confusion matrix for conflict with lead-vehicle events, 0.9 invalid event probability.

		Event classified as:	
		Invalid	**Valid**
Event was:	**Invalid**	**72.2**	27.8
	Valid	23.9	**76.1**

Crashes are still being classified correctly 76.9 percent of the time, but near-crash correct classification now drops to 76.1 percent. However, the percentage of invalid events classified correctly increases to 72.2 percent.

Overall, these analyses suggest that expected probabilities can be used to tradeoff missed valid events with a sufficiently small number of invalid events that will be included in the dataset. This concept was then combined with additional filtering of the data that personalized the analysis for each participant, as discussed below.

The purpose of these filters was to invalidate data before the discriminant analysis process. To limit the scope of the analysis, only conflicts with lead and following vehicles were considered (see rationale above). However, results suggested that the filtering approach hindered the performance of the analyses for conflicts with following vehicles. Thus, results for that particular set of analyses are not presented here; instead, only results for conflict with lead vehicle are shown. In addition, the personalization created by a portion of the filtering process required that only primary drivers were used in the analysis.

The initial filter was applied to the data to attempt to reduce the number of invalid events considered. This filter required any valid triggers to have a minimum Forward TTC of less than 10s, with a maximum range of 50 m (164 ft) and at least one observation when the range rate was negative. This filter was applied to data across participants, and reduced the number of valid events by 36.0 percent and the number of invalid events by 30.1 percent. In raw numbers for the samples used, however, this translates to 131 lost valid events (out of 364) versus 5,100 invalid events that would be correctly ignored (out of 16,927), a large practical difference.

The second filter, which was implemented in addition to the Forward TTC filter described above, attempted to account for some of the individual differences between drivers before the discriminant analysis was performed. This approach assumes that a DAS used in a large-scale study would have some sort of learning capability that would tailor the triggers to each driver. The idea for this approach comes from experimental observations discussed elsewhere in this report. What was normal driving for some drivers represented a critical incident or near-crash for others. This variability increases the noise in a discriminant analysis that is examining aggregate data.

The personalized filter consisted of eliminating from consideration observations that exceeded preset longitudinal deceleration thresholds for each driver. Table 13.19 lists the effects of the filter depending on the deceleration threshold selected. Even when the percentile threshold was set as high as 50 percent, a small number of additional valid events were lost (22 when compared to those lost due to the Forward TTC filter by itself) when compared to the 2,913 additional invalid events that would not have to be reduced.

Table 13.19. Valid events remaining as a function of deceleration threshold selected for pre-discriminant analysis filter.

Deceleration Threshold (Calculated for Each Driver)	Number of Valid Events Evaluated	Percentage of Valid Events Lost (%)	Number of Invalid Events Evaluated	Percentage of Invalid Events Correctly Rejected (%)
10th percentile	231	36.5	10,511	37.9
20th percentile	229	37.1	10,268	39.3
30th percentile	223	38.7	9,941	41.3
40th percentile	217	40.4	9,481	44.0
50th percentile	211	42.0	8,914	47.3

These filtering processes also had, as expected, some positive effects on the results from the discriminant analyses. Table 13.20 combines the 40th and 50th percentiles filters with different values for expected probabilities. Classification accuracies can be improved by manipulating these probabilities, as was shown before. The improvements followed similar patterns for both percentile threshold settings. Results also improved when compared to the pre-filtering numbers, previously presented. While the percentage of correct invalid events tended to be larger than it was for the pre-filtered data, this was mainly due to the large number of invalid events that were removed by the filtering procedure. There are a fewer number of incorrectly classified invalid events for the filtering approach than for the approach that did not filter the data.

Table 13.20. Discriminant analysis classification results as a function of expected probability and percentile threshold.

	50th Percentile Threshold		40th Percentile Threshold	
Probability of Valid/Probability of Invalid	Invalid Classified as Invalid (%)	Valid Classified as Valid (%)	Invalid Classified as Invalid (%)	Valid Classified as Valid (%)
Based on Sample	79.5	72.5	79.8	71.9
0.1 / 0.9	56.8	82.5	58.9	82.5
0.2 / 0.8	46.9	87.2	49.4	86.2
0.3 / 0.7	41.8	88.1	43.7	87.6
0.4 / 0.6	37.7	89.6	39.8	87.6
0.5 / 0.5	34.5	90.5	36.4	90.8

Results for the analyses described in this section point to several conclusions that are relevant for a large-scale naturalistic data collection effort. First, crashes and near-crashes should be the focus of such an effort. Incidents are observed at a much higher rate than crashes and near-crashes; a total of 90.9 percent of all valid events were classified as incidents. Including incidents would likely overwhelm any data reduction effort for a large-scale study. Incidents are also closer in terms of kinematic signature to many invalid events than are crashes and near-crashes, making their discrimination more difficult.

Second, assuming that crashes and near-crashes are the focus of a large-scale study, tradeoffs concerning loss of valid events should focus on losing a minimal number of near-crashes. Based on the results of the discriminant analyses, changes in the sensitivity of the analysis had minimal effects on the number of crashes detected, but affected to a larger extent the number of near-crashes detected. Maximizing the number of near-crashes detected while minimizing the number of invalid events also tends to maximize the number of crashes detected.

Third, it seems that tailoring the triggering algorithms to particular individuals is a feasible partial solution to minimizing the number of invalid triggers collected, when it is combined with appropriately selected expected probabilities. In the analyses above, this process was very effective in reducing the number of invalid events detected. Assuming that the 40th percentile longitudinal acceleration threshold is used along with expected probabilities based on our current sample, only 20.2 percent of invalid events would be kept, compared to 71.9 percent of the valid events, as shown on Table 13.21.

Table 13.21. Confusion matrix for conflict with lead-vehicle events, 40th percentile longitudinal acceleration, probabilities based on sample.

		Event classified as:	
		Invalid	**Valid**
Event was:	**Invalid**	**79.8**	20.2
	Valid	28.1	**71.9**

It would also be expected that the majority of the valid events lost would be near-crashes, rather than crashes, given particular aspects of the crash event severity (e.g., longitudinal acceleration spikes) that make them easy to identify.

Achieving this tailoring process in a large-scale study would require some initial data collection on each participant's driving habits that would then be used to tailor the triggers for that driver, which should always be the primary driver for the vehicle. This data collection period might be as short as a week, based on the data obtained for this study. While a small additional investment would be required to achieve this goal, the benefit gained by shortening the data reduction effort seems attractive.

Table 13.21 represents a reasonable tradeoff between missing valid events and avoiding numerous invalid events. As shown throughout this section, various settings can produce drastic changes in the composition of the table. The percentage of valid events collected can reach 90 percent in some cases (40th percentile threshold, 0.5/0.5 expected probabilities), but this occurs at a considerable cost in terms of the percentage of invalid events that have to be reduced (~60%). Compared to Table 13.21, this implies that an additional 40 percent of invalid events would have to be reduced to gain less than 20 percent for the number of valid events detected. Therefore, tradeoffs can be made, depending on the goal of a large-scale data collection effort. However, provided the large number of cars that would be on the road in such a case (~5,000), losing a relatively small percentage of near-crash events seems a reasonable cost to substantially reduce the data reduction effort. Chapter 14, Goal 10 (separate report) presents a sample data reduction savings calculation based on the percentages suggested in Table 13.21.

Logistic Regression. Logistic regression was explored as an alternative to discriminant analysis. Recall that discriminant analysis has an implicit requirement of multivariate normality, which was not always present within the predictor variables used. Logistic regression relaxes this requirement.

Regardless of this advantage, logistic regression performed poorly on all the prediction trials for which it was used. While it was run on different sets of data, similar to those discussed for discriminant analysis, logistic regression always performed much more poorly than discriminant analysis. For example, for events of a conflict with lead-vehicle nature, considering SUVs and cars, and classifying events as either valid or invalid, the results are summarized in the following confusion matrix:

Table 13.22. Confusion matrix for conflict with lead-vehicle events using the logistic regression approach.

		Event classified as:	
		Invalid	**Valid**
Event was:	**Invalid**	**99.9**	0.1
	Valid	92.7	**7.3**

This result was typical of the logistic regression approach. Invalid events were classified correctly a large percentage of the time, but valid events were predicted correctly a very small percentage of the time (7.3%).

Summary

Overall, using the aggregate sensor data produced higher levels of performance than data from any one sensor. This was not surprising, given that different events had different natures. Each of these event natures had a particular set of kinematic characteristics that best identified it. For example, conflicts with lead vehicles can be expected to exhibit larger longitudinal decelerations than conflicts with side vehicles, when the lateral acceleration is expected to be higher. Any method that aggregates the data from the different sensors should therefore be able to discriminate better than data from a single sensor.

The tradeoff between collecting events of various severities is also important. Incidents occurred at a much higher rate than crashes and near-crashes, and they were not considered in many of the analyses of this section. For a large-scale naturalistic data collection effort, it seems that the number of incidents, if collected, would be higher than necessary and would place an undue burden on any data reduction capabilities. They also exhibit closer kinematic profiles to invalid events, making discrimination much more difficult. This is a lesser problem with near-crashes and a much smaller problem for crashes. As evidence, note the stability in the number of crashes correctly classified by the discriminant analysis, even when expected probabilities for the events were changed. This stability was reduced for near-crashes and would be expected to be further reduced for incidents.

It is important to note that the presence of a crash was typically indicated by a variety of sources for this study and this would also likely be true in the next study. These sources included: subject self-reports, notations from data downloaders, and greater-than 1.0g acceleration spikes

in the data stream. Thus, the classification problem for a large-scale study would primarily be one of correct identification of near-crash events.

Based on the results of the best classification scheme obtained, it seems that the acceleration, forward headway, and yaw rate sensors resulted in the most effective classification performance when considered independently. Their performance is increased by aggregating their information. It seems that a suite of these sensors and the appropriate signal processing (i.e., filtering and combining sensor inputs) could yield a system that collects a relatively high percentage of valid events and a comparatively small percentage of invalid events. Data reduction would still be needed, but reductionists would have to sift through a smaller number of invalid events for each valid event found. This is particularly true when an algorithm can be tuned for an individual driver using their first week of driving data to increase or reduce sensitivity to likely near-crash event signatures.

Based on the initial multivariate discriminant algorithms described above, (which are fairly consistent with the kinematic "minimum error" criteria discussed in Chapter 6, *Goal* 2), one could expect to capture virtually all of the crashes and at least 72 percent of the near-crash events in a large-scale study while rejecting 80 percent of the invalid events. For a 5,000 vehicle, 10,000 driving-year large-scale study, this would result in a database of roughly 2,400 crashes/collisions (of all severities as defined here) and almost 50,000 (out of 65,000 total) near-crash events, excluding any sampling of other incidents, specific circumstances of interest, or baseline events.

DISCUSSION

The DAS used in this study purposefully contained a large number of sensors, some of which were redundant, with the goal of maximizing the level of redundancy within the system and obtaining a dataset that represented a nearly best-case scenario of data availability. This large number of sensors may not be needed for a larger-scale study. The events of interest may be more narrowly targeted or the magnitude of the data large enough that missing a few valid events is not as important as minimizing the number of invalid events that contaminate the dataset and increase the data reduction effort.

A larger-scale study would also magnify any system repair and/or maintenance needs. Thus, reducing the number of sensors and selecting sensors with low associated failure rates would be an important aspect of such an effort. Most of the sensors used in the data collection effort reported herein had very low failure rates, which will likely be even lower as technology progresses. The most problem-prone sensors were video and radar.

Given the advantages of video, however, it seems that its place as a sensor in a larger-scale study is necessary, albeit a smaller number of cameras might be acceptable (see the *Goal 10 Report* for further discussion of this issue). While the performance of the sensors in discriminating between valid and invalid events can be increased by data analysis methods, this increase is not large enough to warrant the elimination of the only method available for event verification.

The failure rate for radar was lower than for video. While there are problems with radar data, the radar units have to be carefully installed, and they are usually damaged in crashes, the relative

position and speed of leading traffic are important factors to consider for triggering to obtain valid events. Thus, despite the failure rate, the technology would be needed for a larger-scale study. Of course, if other technologies could sense the same data with a lower failure rate, they should be considered. At this time, however, no such technology exists at a reasonable price.

Other sensors, including accelerometers and gyros (for yaw rate), had negligible failure rates, undetectable for the current study. These sensors also provided data that proved very useful for valid event discrimination, as discussed in Question 3. These sensors should be included in the sensor suite for a large-scale study.

The triggers used in such an array of sensors would likely take values similar to those discussed in this goal, and the discrimination process using aggregate data would likely be equivalent. However, some of these triggers may become more stringent if higher accuracy sensors are used or if the data collection rate for some of the sensors is increased. Thus, the numbers suggested in this section for future use represent good starting values and their performance should be tested within the final system in which they are included.

CHAPTER 14: GENERAL CONCLUSIONS

The 100-Car Naturalistic Driving Study is the first instrumented vehicle study undertaken with the primary purpose of collecting large-scale naturalistic driving data. Two hundred forty-one participants between the ages of 18 and 73 drove for a total of over 2.1 MVMT and 42,300 hours over an 18-month data collection period. Drivers were given no special instructions, no experimenter was present, and the data collection instrumentation was unobtrusive. In addition, the majority of the drivers drove their own vehicle (78 out of 100 vehicles). As described throughout this document, there is every indication that the drivers rapidly disregarded the presence of the instrumentation. Thus, the resulting database contains many extreme cases of driving behavior and performance including extreme drowsiness, impairment, judgment error, risk taking, willingness to engage in secondary tasks, aggressive driving, and traffic violations, among others, that have been heretofore greatly attenuated by other empirical techniques.

ADVANTAGES TO THE NATURALISTIC APPROACH

Five channels of digitally compressed video and numerous electronic sensors including radars and accelerometers comprised the data collection system. A variety of data reduction and analysis tools were created to allow for efficient utilization of the resulting six-terabyte raw database. Using both the video and electronic sensor data, an "event" database consisting of crash, near-crash and crash-relevant conflict events was created. This database, consisting of almost 10,000 such events, was used to address 10 specific objectives ranging from the factors that contribute to rear-end and lane change events, to the prevalence of distraction and drowsiness in crashes and near-crashes, to precise analysis of the factors leading up to a crash or near-crash event. The specific results and conclusions associated with these 10 objectives are described in Chapters 4 through 13 of this report. In addition, these results are summarized in the Report Overview presented at the beginning of this document.

A real strength of this approach is that the event database can be potentially for years to come, to address a multitude of additional research questions beyond those originally conceptualized. As of this writing, the database is being expanded to answer additional research objectives. This expansion includes the development of a statistically valid baseline to assess exposure for evaluation of findings. This first baseline includes 20,000 random samples of segments of driving behavior.

Another significant advantage of this approach over existing approaches is that the video allows direct viewing of all of the pre-event and during-event parameters, including the pre-event driver behaviors such as distraction, drowsiness, error, and so forth. In addition, this technique allows the precise calculation of parameters such as vehicle speed, vehicle headway, time-to-collision, and driver reaction time.

Naturalistic methods have the potential to fill a void in our existing driving safety research. Specifically, it provides much more detailed and accurate information regarding near-crash, pre-crash, and crash events than is currently available, even after a detailed crash investigation. Police reports and crash investigations rely on eyewitness accounts. Such data have been shown to be limited in accuracy. For example, drivers often do not remember specific details that occur

very rapidly as a crash or near-crash scenario unfolds. This is exacerbated by cases in which the drivers or passengers have been dazed in a crash event, or are trying to hide the details of what occurred due to reasons of embarrassment or fear of prosecution.

Furthermore, the data provides much greater external validity relative to the larger context of driving when compared to empirical methods such as test tracks or simulators. Unlike empirical methods, naturalistic studies allow the consideration of many factors simultaneously including questions such as, "How do drivers modify their risk to situations where they choose to engage in a potentially distracting task?" "Do drivers increase headway, reduce speed, or wait for a straight stretch of road thereby mitigating their crash risk?" "When they do exhibit such adaptation behavior, do drivers tend to over- or under-compensate for a given situation?" These questions cannot be effectively addressed using conventional empirical methods. This has always limited our ability to fully understand the relationship between surrogate measures of safety, such as lane keeping performance or eyeglance behavior, and crash risk. Furthermore, as demonstrated repeatedly in the 100-Car Study, the absence of an experimenter avoids potential modification of drivers' performance and behavior that may occur in contrived empirical circumstances.

For the first time, we are collecting detailed information on large numbers of near-crash events. These events were operationally defined for this study as having the presence of identical elements to a crash scenario, with the exception of the presence of a successful evasive maneuver. These types of events have two important features that crash data do not. First, they occur much more frequently (e.g., 15 times more than crashes). Second, near-crash events are cases where a driver successfully performed an evasive maneuver. Understanding these cases may give additional insight into the factors that allow drivers to be effective defensive drivers, as well as potential countermeasures to aid these drivers.

IMPLICATIONS FOR A LARGE-SCALE EFFORT

Despite the massive scope of the current effort, it was designed to also serve as a pilot to a much larger future study. From an epidemiological viewpoint, the study was small with the presence of 15 police-reported and 82 total crashes and minor collisions. Furthermore, drivers were recruited from only one area of the country (Northern Virginia/Washington, DC, metro area). One purpose of a larger-scale study would be to have a statistically representative sample of crashes (perhaps 2,000) and a more representative subject/environment sample.

Since a primary purpose of the 100-Car Study was to serve as a pilot for a larger-scale (e.g., a 5,000-car) study, a goal was to evaluate the process and results of the 100-Car Study to assess the feasibility of such an undertaking. Based upon the results of the evaluations conducted, it is believed that a large-scale database would be an enormous asset and would be used by transportation researchers for many years. Such an undertaking would allow researchers to gain insight and understanding into a wide array of driving behavior issues and potentially serve as a basis for decision making and program development within both the government and business sectors. This belief is based upon the robustness of these pilot results and the anticipation that these data will continue to be analyzed and the results made available, from a variety of researchers and research organizations for, at least, the next 5 to 10 years. Clearly, these data can provide unique insights into issues that have eluded the highway safety community.

SUMMARY

It is believed that the database that has resulted from this study can be useful for a variety of purposes for a number of years. In addition, the initial event database described above can be continually enhanced, since all of the video and electronic data for the entire study have been archived. The current project specified 10 objectives or *goals* that would be addressed through the initial analysis of the event database. This report addresses these 10 goals. Some of most important findings addressed as part of the analysis of these 10 goals are presented below:

- This study allowed, perhaps for the first time, the capture of crash and collision events that included minor, non-property-damage contact. These low-severity collisions provide very valuable information and occur much more frequently than more severe crashes. As a result, crash/collision-involvement was much higher than expected in that 82 total crashes/collisions were reported in this study, while only 15 of these crashes were reported to the police. For urban/suburban settings, this suggests that total crash/collision involvement may be over five times higher than police-reported crashes.

- Almost 80 percent of all crashes and 65 percent of all near-crashes involved the driver looking away from the forward roadway just prior to the onset of the conflict. Prior estimates related to "distraction" as a contributing factor have been in the range of 25 percent.

- Inattention, which was operationally defined as including: (1) secondary task distraction, (2) driving-related inattention to the forward roadway (e.g., blind spot checks), (3) moderate to extreme drowsiness, and (4) other non-driving-related eyeglances, was a contributing factor for 93 percent of the conflict with lead vehicle crashes and minor collisions. In 86 percent of the lead vehicle crashes/collisions, the headway at the onset of the event was greater than 2.0 seconds.

- For scenarios involving conflict with a lead vehicle, the most frequent cases of lower severity conflicts (i.e., incidents and near-crashes) occurred in lead-vehicle moving scenarios, while 100 percent of the crashes (14 total) occurred when the lead vehicle was stopped. This indicates that drivers have sufficient awareness and ability to perform evasive maneuvers when closing rates are lower and/or expectancies about the flow of traffic are not violated.

- The rate of inattention-related crash and near-crash events decreases dramatically with age, with the rate being as much as four times higher for the 18- to 20-year-old age group relative to some of the older driver groups (i.e., 35 and up).

- The use of hand-held wireless devices (primarily cell phones but including a small amount of PDA use) was associated with the highest frequency of secondary-task distraction-related events. This was true for both events of lower severity (i.e., incidents) and for events of higher severity (i.e., near-crashes). Wireless devices were also among the categories associated with the highest frequencies of crashes and minor collisions,

along with looking at/reaching for an object in vehicle and passenger-related secondary tasks.

- Drowsiness also appears to affect crashes and collisions at much higher rates than is reported using existing crash databases. Drowsiness was a contributing factor in 12 percent of all crashes and 10 percent of near-crashes, while most current database estimates place drowsiness-related crashes at approximately 2 percent to 4 percent of total crashes.

- The lead-vehicle crash and near-crash data clearly show that development of purely quantitative near-crash criteria (i.e., not requiring at least some degree of verification by a human analyst) is not currently feasible. A primary reason for this was that vehicle kinematics associated with near-crashes were virtually identical to common driving situations that were not indicative of crash risk. Thus, qualitative and quantitative criteria are dependent upon one another to some degree. Fortunately, advances in digital video compression and storage technology, and the advancement of data reduction software, have made video verification feasible for large numbers of events.

- Results from the analysis investigating driver adaptation to instrumented vehicles indicates that even when the same driver was switched from a private vehicle to a leased vehicle, there were still more events per mile in the leased vehicle than in the private vehicle. If there was an effect of adaptation, it was extinguished before the first week of driving was completed. In addition, drivers appeared to adapt to the presence of the unobtrusive instrumentation within the first hour of driving.

In addition to the 10 high-priority goals addressed as part of this report, there are 3 additional research contracts in place to perform further data reduction and analysis efforts for the purpose of addressing another 8 goals. There is also considerable interest in using the data for even more purposes from researchers in several disciplines. Progressing toward this potential for a multi-purpose, highly flexible and adaptable tool for driving safety may be the most important aspect of this study.

REFERENCES

Campbell, B.N., Smith, J.D., and Najm, W.G. (2003). *Examination of Crash Contributing Factors Using National Crash Databases.* (Technical Report No. DOT HS 809 664). National Highway Traffic Safety Administration: Washington, DC.

Cicchetti, D.V., and Feinstein A.R. (1990a). "*High agreement but low kappa: I. The Problems of Two Paradoxes.*" Journal of Clinical Epidemiology, 43, 543-549.

Cicchetti, D.V., and Feinstein A.R. (1990b). "*High agreement but low kappa: II. Resolving the paradoxes.*" Journal of Clinical Epidemiology, 43, 551-558.

Dingus, T.A. (1997). *Empirical data needs in support of model building and benefits estimation.* ITS America (Ed.), Modeling ITS collision avoidance system benefits: expert panel proceedings. Washington, DC: Intelligent Transportation Society of America.

Dingus, T.A., McGehee, D., Hulse, M., Jahns, S., Manakkal, N., Mollenhauer, M., and Fleischman, R. (1995). *TravTek evaluation task C3 camera car study.* (Technical Contract Report No. FHWA-RD-94-076). Washington, DC: Federal Highway Administration.

Dingus, T.A., Neale, V.L., Garness, S.A., Hanowski, R.J., Keisler, A.S., Lee, S.E., Perez, M.A., Robinson, G.S., Belz, S.M., Casali, J.G., Pace-Schott, E.F., Stickgold, R.A., and Hobson, J.A. (2001). *Impact of sleeper berth usage on driver fatigue.* (Technical Contract Report No. DTFH61-96-00068). Washington, DC: Federal Motor Carrier Safety Administration.

Greenberg, R. S., Daniels, S. R., Flanders, W. D., Eley, J. W., and Boring, J. R. (1993). *Medical Epidemiology*. Norwalk, CT: Appleton & Lange.

Hanowski, R.J., Wierwille, W.W., Garness, S.A., and Dingus, T.A. (2000). *Impact of local/short haul operations on driver fatigue.* (Technical Contract Report No. DTFH61-96-C-00105). Washington, DC: Federal Motor Carrier Safety Administration.

Heinrich, H. W., Petersen, D., and Roos, N. (1980). *Industrial Accident Prevention.* New York: McGraw-Hill Book Company.

Jenness, J. W., Lattanzio, R. J., O'Toole, M., Taylor, N., and Pax, C. (2002). *Effects of manual versus voice-activated dialing during simulated driving.* Perceptual and Motor Skills, 94, p. 363-379.

Kiefer, R.J., Cassar, M.T., Flannagan, C.A., LeBlanc, D.J., Palmer, M.D., Deering, R.K., and Shulman, M. (2003). *Forward Collision Warning Requirements Project Final Report Task 1: Refining the CAMP Crash Alert Timing Approach by Examining "Last-Second" Braking and Lane-Change Maneuvers Under Various Kinematic Conditions.* (DOT HS 809 574). National Highway Transportation Safety Administration. Washington, DC.

Knipling, R.R., Wang, J.S., and Yin. S.M. (1993). *Rear-end crashes: Problem size assessment and statistical description* (NHTSA Technical Report No. DTNH-22-91-C-03131). Washington, DC: National Highway Traffic Safety Administration.

Mollenhauer, M. (1998). *Proactive driving safety evaluation: An evaluation of an automated traveler information system and investigation of hazard analysis data*. Unpublished doctoral dissertation, University of Iowa, Iowa City, IA.

Najm, W.G. with Sen. B, Smith, J.D., and Campbell, B.N. (2003). *Analysis of Light Vehicle Crashes and Pre-Crash Scenarios Based on the 2000 General Estimates System*. DOT VNTSC NHTSA 02 04, DOT HS 809 573.

National Highway Traffic Safety Administration. (2002). *Traffic safety facts*. Chicago, IL: National Safety Council.

Neale, V.L., Klauer, S.G., Knipling, R.R., Dingus, T.A., Holbrook, G.T., and Petersen, A. (2002). *The 100-Car Naturalistic Driving Study, Phase 1 – Experimental Design.* (DOT Report HS 809 536). Washington, DC: National Highway Traffic Safety Administration.

Smith, D.L., Najm, W.G., and Lam, A.H. (2003). *Analysis of braking and steering performance in car-following scenarios.* (SAE Technical Paper No. 2003-01-0283). Warrendale, PA: Society of Automotive Engineers.

Smith, D.L., Najm, W.G., Lam, A.H. (2003). *Analysis of Braking and Steering Performance in Car-Following Scenarios.* (SAE Technical Paper Series No. 2003-01-0283). 2003 SAE World Congress, Detroit, MI.

Stutts, J., Feaganes, J., Rodgman, E., Hamlett, C., Meadows, T., Reinfurt, D., Gish, K., Mercadante, M., and Staplin, L. (2003). *Distraction in everyday driving.* Washington, DC: AAA Foundation for Traffic Safety.

Tijerina, L., (2004). *Issues in the evaluation of driver distraction associated with in-vehicle information and telecommunications systems* (www-nrd.nhtsa.dot.gov/departments/nrd-13/driver distraction/pdf/3.pdf)

Wang, J.S., Knipling, R.R., and Goodman, M.J. (1996). *The role of driver inattention in crashes: New statistics from the 1995 Crashworthiness Data System*. 40th Annual Proceedings of the Association for the Advancement of Automotive Medicine, Vancouver, Canada.

Wierwille, W. W. and Ellsworth, L. A. (1994). *Evaluation of driver fatigue by trained raters.* Accident Analysis and Prevention, *26*(4), 571-578.

Wierwille, W. W. and Ellsworth, L. A. (1994). *Evaluation of driver fatigue by trained raters.* Accident Analysis and Prevention, *26*(4), 571-578.

Wierwille, W. W., Hanowski, R. J., Hankey, J. M., Kieliszewski, C. A., Lee, S. E., Medina, A., Keisler, A. S., and Dingus, T. A. (2001) *Identification of driver errors: overview and recommendations.* (Technical Contract Report No. FHWA-RD-02-003). Washington, DC: Federal Highway Administration.

DOT HS 810 593
April 2006

U.S. Department
of Transportation
**National Highway
Traffic Safety
Administration**

www.nhtsa.dot.gov
nhtsa
people saving people